Attorney Docket No.: 018623-014400US
Client Reference No.: EPI 0144.00US

PATENT APPLICATION

INDUCING CELLULAR IMMUNE RESPONSES TO CARCINOEMBRYONIC ANTIGEN USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

Inventor(s): John Fikes, a United States citizen, residing at
6494 Lipmann Street
San Diego, California 92122

Alessandro Sette, an Italian citizen, residing at
5551 Linda Rosa Avenue
La Jolla, California 92037

John Sidney, a United States citizen, residing at
4218 Corte de la Siena
San Diego, California 92130

Scott Southwood, a United States citizen, residing at
10679 Strathmore Drive
Santee, California 92071

Robert Chesnut, a United States citizen, residing at
1473 Kings Cross Drive
Cardiff-by-the-Sea, California 92007

Esteban Celis, a United States citizen, residing at
3683 Wright Road S.W.
Rochester, Minnesota 55902

Elissa Keogh, a United States citizen, residing at
4343 Caminito del Diamante
San Diego, California 92121

TOWNSEND and TOWNSEND and CREW LLP
Two Embarcadero Center, 8th Floor
San Francisco, California 94111-3834
(415) 576-0200

PATENT

Attorney Docket No.: 018623-014400US

INDUCING CELLULAR IMMUNE RESPONSES TO CARCINOEMBRYONIC ANTIGEN USING PEPTIDE AND NUCLEIC ACID COMPOSITIONS

CROSS-REFERENCES TO RELATED APPLICATIONS

This application is a Continuation-In-Part ("CIP") of U.S.S.N. 09/189,702, filed 11/10/98, which is a CIP of U.S.S.N 08/205,713 filed 3/4/94, which is a CIP of abandoned U.S.S.N. 08/159,184 filed 11/29/93, which is a CIP of abandoned U.S.S.N. 08/073,205 filed 6/4/93 which is a CIP of abandoned U.S.S.N 08/027,146 filed 3/5/93. The present application is also related to U.S.S.N. 09/226,775, which is a CIP of abandoned U.S.S.N. 08/815,396, which claims benefit of abandoned U.S.S.N. 60/013,113. Furthermore, the present application is related to U.S.S.N. 09/017,735, which is a CIP of abandoned U.S.S.N. 08/589,108; U.S.S.N. 08/454,033; and U.S.S.N. 08/349,177. The present application is also related to U.S.S.N. 09/017,524, U.S.S.N. 08/821,739, which claims benefit of abandoned U.S.S.N. 60/013,833; and U.S.S.N. 08/347,610, which is a CIP of U.S.S.N. 08/159,339, which is a CIP of abandoned U.S.S.N. 08/103,396, which is a CIP of abandoned U.S.S.N. 08/027,746, which is a CIP of abandoned U.S.S.N. 07/926,666. The present application is also related to U.S.S.N. 09/017,743, which is a CIP of abandoned U.S.S.N. 08/590,298; and U.S.S.N. 08/452,843, which is a CIP of U.S.S.N. 08/344,824, which is a CIP of abandoned U.S.S.N. 08/278,634. The present application is also related to PCT application 99/12066 filed 5/28/99 which claims benefit of provisional U.S.S.N. 60/087,192, and U.S.S.N. 09/009,953, which is a CIP of abandoned U.S.S.N. 60/036,713 and abandoned U.S.S.N. 60/037,432. In addition, the present application is related to U.S.S.N. 09/098,584, U.S.S.N. 09/239,043, U.S.S.N. 60/117,486, U.S.S.N. 09/350,401, and U.S.S.N. 09/357,737. In addition, the present application is related to U.S. Patent Application entitled "Inducing Cellular Immune Responses to p53 Using Peptide and Nucleic Acid Compositions", Attorney Docket No. 018623-014500, filed of even date herewith; U.S. Patent Application entitled "Inducing Cellular Immune Responses to MAGE2/3 Using Peptide and Nucleic Acid Compositions", Attorney Docket No. 018623-014600, filed of even date herewith; and U.S. Patent Application entitled "Inducing Cellular Immune Responses to HER2/neu Using Peptide and Nucleic Acid Compositions", Attorney Docket No. 018623-014800, filed of even date herewith. All of the above applications are incorporated herein by reference.

FEDERALLY SPONSORED RESEARCH AND DEVELOPMENT

This invention was funded, in part, by the United States government under grants with the National Institutes of Health. The U.S. government has certain rights in this invention.

INDEX

I. BACKGROUND OF THE INVENTION

A growing body of evidence suggests that cytotoxic T lymphocytes (CTL) are important in the immune response to tumor cells. CTL recognize peptide epitopes in the context of HLA class I molecules that are expressed on the surface of almost all nucleated cells. Following intracellular processing of endogenously synthesized tumor antigens, antigen-derived peptide epitopes bind to class I HLA molecules in the endoplasmic reticulum, and the resulting complex is then transported to the cell surface. CTL recognize the peptide-HLA class I complex, which then results in the destruction of the cell bearing the HLA-peptide complex directly by the CTL and/or via the activation of non-destructive mechanisms, *e.g.*, activation of lymphokines such as tumor necrosis factor-α (TNF- α) or interferon-γ (IFNγ) which enhance the immune response and facilitate the destruction of the tumor cell.

Tumor-specific helper T lymphocytes (HTLs) are also known to be important for maintaining effective antitumor immunity. Their role in antitumor immunity has been demonstrated in animal models in which these cells not only serve to provide help for induction of CTL and antibody responses, but also provide effector functions, which are mediated by direct cell contact and also by secretion of lymphokines (*e.g.*, IFNγ and TNF- α).

A fundamental challenge in the development of an efficacious tumor vaccine is immune suppression or tolerance that can occur. There is therefore a need to establish vaccine embodiments that elicit immune responses of sufficient breadth and vigor to prevent progression and/or clear the tumor.

The epitope approach, as we have described, may represent a solution to this challenge, in that it allows the incorporation of various antibody, CTL and HTL epitopes, from discrete regions of a target TAA, and/or regions of other TAAs, in a single vaccine composition. Such a composition may simultaneously target multiple dominant and subdominant epitopes and thereby be used to achieve effective immunization in a diverse population.

Carcinoembryonic antigen (CEA) is a 180 kD cell surface and secreted glycoprotein overexpressed on most human adenocarcinomas including colon, rectal, pancreatic and gastric (Muraro *et al.*, *Cancer Res.* 45:5769-5780, 1985) as well as 50% of breast (Steward *et al.*, *Cancer (Phila)* 33:1246-1252, 1974) and 70% of non-small cell lung carcinomas (Vincent *et al.*, *J. Thorac. Cardiovasc. Surg.* 66:320-328, 1978). CEA is

also expressed, to some extent, on normal epithelium and in some fetal tissues (Thompson *et al.*, *J. Clin. Lab. Anal.* 5:344-366, 1991). The abnormally high expression on cancer cells makes CEA an important target for immunotherapy.

The information provided in this section is intended to disclose the presently understood state of the art as of the filing date of the present application. Information is included in this section which was generated subsequent to the priority date of this application. Accordingly, information in this section is not intended, in any way, to delineate the priority date for the invention.

II. SUMMARY OF THE INVENTION

This invention applies our knowledge of the mechanisms by which antigen is recognized by T cells, for example, to develop epitope-based vaccines directed towards TAAs. More specifically, this application communicates our discovery of specific epitope pharmaceutical compositions and methods of use in the prevention and treatment of cancer.

Upon development of appropriate technology, the use of epitope-based vaccines has several advantages over current vaccines, particularly when compared to the use of whole antigens in vaccine compositions. For example, immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-based vaccines. Such immunosuppressive epitopes may, *e.g.*, correspond to immunodominant epitopes in whole antigens, which may be avoided by selecting peptide epitopes from non-dominant regions (*see, e.g.*, Disis *et al.*, *J. Immunol.* 156:3151-3158, 1996).

An additional advantage of an epitope-based vaccine approach is the ability to combine selected epitopes (CTL and HTL), and further, to modify the composition of the epitopes, achieving, for example, enhanced immunogenicity. Accordingly, the immune response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches.

Another major benefit of epitope-based immune-stimulating vaccines is their safety. The possible pathological side effects caused by infectious agents or whole protein antigens, which might have their own intrinsic biological activity, is eliminated.

An epitope-based vaccine also provides the ability to direct and focus an immune response to multiple selected antigens from the same pathogen (a "pathogen" may be an infectious agent or a tumor-associated molecule). Thus, patient-by-patient variability in

the immune response to a particular pathogen may be alleviated by inclusion of epitopes from multiple antigens from the pathogen in a vaccine composition.

Furthermore, an epitope-based anti-tumor vaccine also provides the opportunity to combine epitopes derived from multiple tumor-associated molecules. This capability can therefore address the problem of tumor-to tumor variability that arises when developing a broadly targeted anti-tumor vaccine for a given tumor type and can also reduce the likelihood of tumor escape due to antigen loss. For example, a breast cancer tumor in one patient may express a target TAA that differs from a breast cancer tumor in another patient. Epitopes derived from multiple TAAs can be included in a polyepitopic vaccine that will target both breast cancer tumors.

One of the most formidable obstacles to the development of broadly efficacious epitope-based immunotherapeutics, however, has been the extreme polymorphism of HLA molecules. To date, effective non-genetically biased coverage of a population has been a task of considerable complexity; such coverage has required that epitopes be used that are specific for HLA molecules corresponding to each individual HLA allele. Impractically large numbers of epitopes would therefore have to be used in order to cover ethnically diverse populations. Thus, there has existed a need for peptide epitopes that are bound by multiple HLA antigen molecules for use in epitope-based vaccines. The greater the number of HLA antigen molecules bound, the greater the breadth of population coverage by the vaccine.

Furthermore, as described herein in greater detail, a need has existed to modulate peptide binding properties, *e.g.*, so that peptides that are able to bind to multiple HLA molecules do so with an affinity that will stimulate an immune response. Identification of epitopes restricted by more than one HLA allele at an affinity that correlates with immunogenicity is important to provide thorough population coverage, and to allow the elicitation of responses of sufficient vigor to prevent or clear an infection in a diverse segment of the population. Such a response can also target a broad array of epitopes. The technology disclosed herein provides for such favored immune responses.

In a preferred embodiment, epitopes for inclusion in vaccine compositions of the invention are selected by a process whereby protein sequences of known antigens are evaluated for the presence of motif or supermotif-bearing epitopes. Peptides corresponding to a motif- or supermotif-bearing epitope are then synthesized and tested for the ability to bind to the HLA molecule that recognizes the selected motif. Those peptides that bind at an intermediate or high affinity *i.e.*, an IC_{50} (or a K_D value) of 500

nM or less for HLA class I molecules or an IC_{50} of 1000 nM or less for HLA class II molecules, are further evaluated for their ability to induce a CTL or HTL response. Immunogenic peptide epitopes are selected for inclusion in vaccine compositions.

Supermotif-bearing peptides may additionally be tested for the ability to bind to multiple alleles within the HLA supertype family. Moreover, peptide epitopes may be analogued to modify binding affinity and/or the ability to bind to multiple alleles within an HLA supertype.

The invention also includes embodiments comprising methods for monitoring or evaluating an immune response to a TAA in a patient having a known HLA-type. Such methods comprise incubating a T lymphocyte sample from the patient with a peptide composition comprising a TAA epitope that has an amino acid sequence described in Tables VII to Table XX or Table XXII which binds the product of at least one HLA allele present in the patient, and detecting for the presence of a T lymphocyte that binds to the peptide. A CTL peptide epitope may, for example, be used as a component of a tetrameric complex for this type of analysis.

An alternative modality for defining the peptide epitopes in accordance with the invention is to recite the physical properties, such as length; primary structure; or charge, which are correlated with binding to a particular allele-specific HLA molecule or group of allele-specific HLA molecules. A further modality for defining peptide epitopes is to recite the physical properties of an HLA binding pocket, or properties shared by several allele-specific HLA binding pockets (*e.g.* pocket configuration and charge distribution) and reciting that the peptide epitope fits and binds to the pocket or pockets.

As will be apparent from the discussion below, other methods and embodiments are also contemplated. Further, novel synthetic peptides produced by any of the methods described herein are also part of the invention.

III. BRIEF DESCRIPTION OF THE FIGURES

not applicable

IV. DETAILED DESCRIPTION OF THE INVENTION

The peptide epitopes and corresponding nucleic acid compositions of the present invention are useful for stimulating an immune response to a TAA by stimulating the production of CTL or HTL responses. The peptide epitopes, which are derived directly or indirectly from native TAA protein amino acid sequences, are able to bind to HLA

molecules and stimulate an immune response to the TAA. The complete sequence of the TAA proteins to be analyzed can be obtained from GenBank. Peptide epitopes and analogs thereof can also be readily determined from sequence information that may subsequently be discovered for heretofore unknown variants of particular TAAs, as will be clear from the disclosure provided below.

A list of target TAA includes, but is not limited to, the following antigens: MAGE 1, MAGE 2, MAGE 3, MAGE-11, MAGE-A10, BAGE, GAGE, RAGE, MAGE-C1, LAGE-1, CAG-3, DAM, MUC1, MUC2, MUC18, NY-ESO-1, MUM-1, CDK4, BRCA2, NY-LU-1, NY-LU-7, NY-LU-12, CASP8, RAS, KIAA-2-5, SCCs, p53, p73, CEA, Her 2/neu, Melan-A, gp100, tyrosinase, TRP2, gp75/TRP1, kallikrein, PSM, PAP, PSA, PT1-1, B-catenin, PRAME, Telomerase, FAK, cyclin D1 protein, NOEY2, EGF-R, SART-1, CAPB, HPVE7, p15, Folate receptor CDC27, PAGE-1, and PAGE-4.

The peptide epitopes of the invention have been identified in a number of ways, as will be discussed below. Also discussed in greater detail is that analog peptides have been derived and the binding activity for HLA molecules modulated by modifying specific amino acid residues to create peptide analogs exhibiting altered immunogenicity. Further, the present invention provides compositions and combinations of compositions that enable epitope-based vaccines that are capable of interacting with HLA molecules encoded by various genetic alleles to provide broader population coverage than prior vaccines.

IV.A. Definitions

The invention can be better understood with reference to the following definitions, which are listed alphabetically:

A "computer" or "computer system" generally includes: a processor; at least one information storage/retrieval apparatus such as, for example, a hard drive, a disk drive or a tape drive; at least one input apparatus such as, for example, a keyboard, a mouse, a touch screen, or a microphone; and display structure. Additionally, the computer may include a communication channel in communication with a network. Such a computer may include more or less than what is listed above.

"Cross-reactive binding" indicates that a peptide is bound by more than one HLA molecule; a synonym is degenerate binding.

A "cryptic epitope" elicits a response by immunization with an isolated peptide, but the response is not cross-reactive *in vitro* when intact whole protein which comprises the epitope is used as an antigen.

A "dominant epitope" is an epitope that induces an immune response upon immunization with a whole native antigen (see, *e.g.*, Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993). Such a response is cross-reactive *in vitro* with an isolated peptide epitope.

With regard to a particular amino acid sequence, an "epitope" is a set of amino acid residues which is involved in recognition by a particular immunoglobulin, or in the context of T cells, those residues necessary for recognition by T cell receptor proteins and/or Major Histocompatibility Complex (MHC) receptors. In an immune system setting, *in vivo* or *in vitro*, an epitope is the collective features of a molecule, such as primary, secondary and tertiary peptide structure, and charge, that together form a site recognized by an immunoglobulin, T cell receptor or HLA molecule. Throughout this disclosure epitope and peptide are often used interchangeably. It is to be appreciated, however, that isolated or purified protein or peptide molecules larger than and comprising an epitope of the invention are still within the bounds of the invention.

"Human Leukocyte Antigen" or "HLA" is a human class I or class II Major Histocompatibility Complex (MHC) protein (*see, e.g.*, Stites, *et al.*, IMMUNOLOGY, 8TH ED., Lange Publishing, Los Altos, CA, 1994).

An "HLA supertype or family", as used herein, describes sets of HLA molecules grouped on the basis of shared peptide-binding specificities. HLA class I molecules that share somewhat similar binding affinity for peptides bearing certain amino acid motifs are grouped into HLA supertypes. The terms HLA superfamily, HLA supertype family, HLA family, and HLA xx-like molecules (where xx denotes a particular HLA type), are synonyms.

Throughout this disclosure, results are expressed in terms of "IC_{50}'s." IC_{50} is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (*i.e.*, limiting HLA proteins and labeled peptide concentrations), these values approximate K_D values. Assays for determining binding are described in detail, *e.g.*, in PCT publications WO 94/20127 and WO 94/03205. It should be noted that IC_{50} values can change, often dramatically, if the assay conditions are varied, and depending on the particular reagents

used (*e.g.*, HLA preparation, *etc.*). For example, excessive concentrations of HLA molecules will increase the apparent measured IC_{50} of a given ligand.

Alternatively, binding is expressed relative to a reference peptide. Although as a particular assay becomes more, or less, sensitive, the IC_{50}'s of the peptides tested may change somewhat, the binding relative to the reference peptide will not significantly change. For example, in an assay run under conditions such that the IC_{50} of the reference peptide increases 10-fold, the IC_{50} values of the test peptides will also shift approximately 10-fold. Therefore, to avoid ambiguities, the assessment of whether a peptide is a good, intermediate, weak, or negative binder is generally based on its IC_{50}, relative to the IC_{50} of a standard peptide.

Binding may also be determined using other assay systems including those using: live cells (*e.g.*, Ceppellini *et al.*, *Nature* 339:392, 1989; Christnick *et al.*, *Nature* 352:67, 1991; Busch *et al.*, *Int. Immunol.* 2:443, 19990; Hill *et al.*, *J. Immunol.* 147:189, 1991; del Guercio *et al.*, *J. Immunol.* 154:685, 1995), cell free systems using detergent lysates (*e.g.*, Cerundolo *et al.*, *J. Immunol.* 21:2069, 1991), immobilized purified MHC (*e.g.*, Hill *et al.*, *J. Immunol.* 152, 2890, 1994; Marshall *et al.*, *J. Immunol.* 152:4946, 1994), ELISA systems (*e.g.*, Reay *et al.*, *EMBO J.* 11:2829, 1992), surface plasmon resonance (*e.g.*, Khilko *et al.*, *J. Biol. Chem.* 268:15425, 1993); high flux soluble phase assays (Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994), and measurement of class I MHC stabilization or assembly (*e.g.*, Ljunggren *et al.*, *Nature* 346:476, 1990; Schumacher *et al.*, *Cell* 62:563, 1990; Townsend *et al.*, *Cell* 62:285, 1990; Parker *et al.*, *J. Immunol.* 149:1896, 1992).

As used herein, "high affinity" with respect to HLA class I molecules is defined as binding with an IC_{50}, or K_D value, of 50 nM or less; "intermediate affinity" is binding with an IC_{50} or K_D value of between about 50 and about 500 nM. "High affinity" with respect to binding to HLA class II molecules is defined as binding with an IC_{50} or K_D value of 100 nM or less; "intermediate affinity" is binding with an IC_{50} or K_D value of between about 100 and about 1000 nM.

The terms "identical" or percent "identity," in the context of two or more peptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using a sequence comparison algorithm or by manual alignment and visual inspection.

An "immunogenic peptide" or "peptide epitope" is a peptide that comprises an allele-specific motif or supermotif such that the peptide will bind an HLA molecule and

induce a CTL and/or HTL response. Thus, immunogenic peptides of the invention are capable of binding to an appropriate HLA molecule and thereafter inducing a cytotoxic T cell response, or a helper T cell response, to the antigen from which the immunogenic peptide is derived.

The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention preferably do not contain materials normally associated with the peptides in their *in situ* environment.

"Major Histocompatibility Complex" or "MHC" is a cluster of genes that plays a role in control of the cellular interactions responsible for physiologic immune responses. In humans, the MHC complex is also known as the HLA complex. For a detailed description of the MHC and HLA complexes, see, Paul, FUNDAMENTAL IMMUNOLOGY, 3RD ED., Raven Press, New York, 1993.

The term "motif" refers to the pattern of residues in a peptide of defined length, usually a peptide of from about 8 to about 13 amino acids for a class I HLA motif and from about 6 to about 25 amino acids for a class II HLA motif, which is recognized by a particular HLA molecule. Peptide motifs are typically different for each protein encoded by each human HLA allele and differ in the pattern of the primary and secondary anchor residues.

A "negative binding residue" or "deleterious residue" is an amino acid which, if present at certain positions (typically not primary anchor positions) in a peptide epitope, results in decreased binding affinity of the peptide for the peptide's corresponding HLA molecule.

The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other, typically by peptide bonds between the α-amino and carboxyl groups of adjacent amino acids. The preferred CTL-inducing peptides of the invention are 13 residues or less in length and usually consist of between about 8 and about 11 residues, preferably 9 or 10 residues. The preferred HTL-inducing oligopeptides are less than about 50 residues in length and usually consist of between about 6 and about 30 residues, more usually between about 12 and 25, and often between about 15 and 20 residues.

"Pharmaceutically acceptable" refers to a non-toxic, inert, and/or physiologically compatible composition.

A "primary anchor residue" is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic peptide and the HLA molecule. One to three, usually two, primary anchor residues within a peptide of defined length generally defines a "motif" for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding grooves of an HLA molecule, with their side chains buried in specific pockets of the binding grooves themselves. In one embodiment, for example, the primary anchor residues are located at position 2 (from the amino terminal position) and at the carboxyl terminal position of a 9-residue peptide epitope in accordance with the invention. The primary anchor positions for each motif and supermotif are set forth in Table 1. For example, analog peptides can be created by altering the presence or absence of particular residues in these primary anchor positions. Such analogs are used to modulate the binding affinity of a peptide comprising a particular motif or supermotif.

"Promiscuous recognition" is where a distinct peptide is recognized by the same T cell clone in the context of various HLA molecules. Promiscuous recognition or binding is synonymous with cross-reactive binding.

A "protective immune response" or "therapeutic immune response" refers to a CTL and/or an HTL response to an antigen derived from an infectious agent or a tumor antigen, which prevents or at least partially arrests disease symptoms or progression. The immune response may also include an antibody response which has been facilitated by the stimulation of helper T cells.

The term "residue" refers to an amino acid or amino acid mimetic incorporated into an oligopeptide by an amide bond or amide bond mimetic.

A "secondary anchor residue" is an amino acid at a position other than a primary anchor position in a peptide which may influence peptide binding. A secondary anchor residue occurs at a significantly higher frequency amongst bound peptides than would be expected by random distribution of amino acids at one position. The secondary anchor residues are said to occur at "secondary anchor positions." A secondary anchor residue can be identified as a residue which is present at a higher frequency among high or intermediate affinity binding peptides, or a residue otherwise associated with high or intermediate affinity binding. For example, analog peptides can be created by altering the presence or absence of particular residues in these secondary anchor positions. Such analogs are used to finely modulate the binding affinity of a peptide comprising a particular motif or supermotif.

A "subdominant epitope" is an epitope which evokes little or no response upon immunization with whole antigens which comprise the epitope, but for which a response can be obtained by immunization with an isolated peptide, and this response (unlike the case of cryptic epitopes) is detected when whole protein is used to recall the response *in vitro* or *in vivo*.

A "supermotif" is a peptide binding specificity shared by HLA molecules encoded by two or more HLA alleles. Preferably, a supermotif-bearing peptide is recognized with high or intermediate affinity (as defined herein) by two or more HLA molecules.

"Synthetic peptide" refers to a peptide that is not naturally occurring, but is man-made using such methods as chemical synthesis or recombinant DNA technology.

The nomenclature used to describe peptide compounds follows the conventional practice wherein the amino group is presented to the left (the N-terminus) and the carboxyl group to the right (the C-terminus) of each amino acid residue. When amino acid residue positions are referred to in a peptide epitope they are numbered in an amino to carboxyl direction with position one being the position closest to the amino terminal end of the epitope, or the peptide or protein of which it may be a part. In the formulae representing selected specific embodiments of the present invention, the amino- and carboxyl-terminal groups, although not specifically shown, are in the form they would assume at physiologic pH values, unless otherwise specified. In the amino acid structure formulae, each residue is generally represented by standard three letter or single letter designations. The L-form of an amino acid residue is represented by a capital single letter or a capital first letter of a three-letter symbol, and the D-form for those amino acids having D-forms is represented by a lower case single letter or a lower case three letter symbol. Glycine has no asymmetric carbon atom and is simply referred to as "Gly" or G. Symbols for the amino acids are shown below.

Single Letter Symbol	Three Letter Symbol	Amino Acids
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

IV.B. Stimulation of CTL and HTL responses

The mechanism by which T cells recognize antigens has been delineated during the past ten years. Based on our understanding of the immune system we have developed efficacious peptide epitope vaccine compositions that can induce a therapeutic or prophylactic immune response to a TAA in a broad population. For an understanding of the value and efficacy of the claimed compositions, a brief review of immunology-related technology is provided.

A complex of an HLA molecule and a peptidic antigen acts as the ligand recognized by HLA-restricted T cells (Buus, S. *et al.*, *Cell* 47:1071, 1986; Babbitt, B. P. *et al.*, *Nature* 317:359, 1985; Townsend, A. and Bodmer, H., *Annu. Rev. Immunol.* 7:601,

1989; Germain, R. N., *Annu. Rev. Immunol.* 11:403, 1993). Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues that correspond to motifs required for specific binding to HLA antigen molecules have been identified and are described herein and are set forth in Tables I, II, and III (*see* also, *e.g.*, Southwood, *et al.*, *J. Immunol.* 160:3363, 1998; Rammensee, *et al.*, *Immunogenetics* 41:178, 1995; Rammensee *et al.*, SYFPEITHI, access via web at : http://134.2.96.221/scripts.hlaserver.dll/home.htm; Sette, A. and Sidney, J. *Curr. Opin. Immunol.* 10:478, 1998; Engelhard, V. H., *Curr. Opin. Immunol.* 6:13, 1994; Sette, A. and Grey, H. M., *Curr. Opin. Immunol.* 4:79, 1992; Sinigaglia, F. and Hammer, J. *Curr. Biol.* 6:52, 1994; Ruppert *et al.*, *Cell* 74:929-937, 1993; Kondo *et al.*, *J. Immunol.* 155:4307-4312, 1995; Sidney *et al.*, *J. Immunol.* 157:3480-3490, 1996; Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; Sette, A. and Sidney, J. *Immunogenetics*, in press, 1999).

Furthermore, x-ray crystallographic analysis of HLA-peptide complexes has revealed pockets within the peptide binding cleft of HLA molecules which accommodate, in an allele-specific mode, residues borne by peptide ligands; these residues in turn determine the HLA binding capacity of the peptides in which they are present. (*See*, *e.g.*, Madden, D.R. *Annu. Rev. Immunol.* 13:587, 1995; Smith, *et al.*, *Immunity* 4:203, 1996; Fremont *et al.*, *Immunity* 8:305, 1998; Stern *et al.*, *Structure* 2:245, 1994; Jones, E.Y. *Curr. Opin. Immunol.* 9:75, 1997; Brown, J. H. *et al.*, *Nature* 364:33, 1993; Guo, H. C. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:8053, 1993; Guo, H. C. *et al.*, *Nature* 360:364, 1992; Silver, M. L. *et al.*, *Nature* 360:367, 1992; Matsumura, M. *et al.*, *Science* 257:927, 1992; Madden *et al.*, *Cell* 70:1035, 1992; Fremont, D. H. *et al.*, *Science* 257:919, 1992; Saper, M. A. , Bjorkman, P. J. and Wiley, D. C., *J. Mol. Biol.* 219:277, 1991.)

Accordingly, the definition of class I and class II allele-specific HLA binding motifs, or class I or class II supermotifs allows identification of regions within a protein that have the potential of binding particular HLA molecules.

The present inventors have found that the correlation of binding affinity with immunogenicity, which is disclosed herein, is an important factor to be considered when evaluating candidate peptides. Thus, by a combination of motif searches and HLA-peptide binding assays, candidates for epitope-based vaccines have been identified. After determining their binding affinity, additional confirmatory work can be performed to select, amongst these vaccine candidates, epitopes with preferred characteristics in terms of population coverage, antigenicity, and immunogenicity.

Various strategies can be utilized to evaluate immunogenicity, including:

1) Evaluation of primary T cell cultures from normal individuals (*see, e.g.*, Wentworth, P. A. *et al.*, *Mol. Immunol.* 32:603, 1995; Celis, E. *et al.*, *Proc. Natl. Acad. Sci. USA* 91:2105, 1994; Tsai, V. *et al.*, *J. Immunol.* 158:1796, 1997; Kawashima, I. *et al.*, *Human Immunol.* 59:1, 1998); This procedure involves the stimulation of peripheral blood lymphocytes (PBL) from normal subjects with a test peptide in the presence of antigen presenting cells *in vitro* over a period of several weeks. T cells specific for the peptide become activated during this time and are detected using, *e.g.*, a ^{51}Cr-release assay involving peptide sensitized target cells.

2) Immunization of HLA transgenic mice (*see, e.g.*, Wentworth, P. A. *et al.*, *J. Immunol.* 26:97, 1996; Wentworth, P. A. *et al.*, *Int. Immunol.* 8:651, 1996; Alexander, J. *et al.*, *J. Immunol.* 159:4753, 1997); In this method, peptides in incomplete Freund's adjuvant are administered subcutaneously to HLA transgenic mice. Several weeks following immunization, splenocytes are removed and cultured *in vitro* in the presence of test peptide for approximately one week. Peptide-specific T cells are detected using, *e.g.*, a ^{51}Cr-release assay involving peptide sensitized target cells and target cells expressing endogenously generated antigen.

3) Demonstration of recall T cell responses from patients who have been effectively vaccinated or who have a tumor; (*see, e.g.*, Rehermann, B. *et al.*, *J. Exp. Med.* 181:1047, 1995; Doolan, D. L. *et al.*, *Immunity* 7:97, 1997; Bertoni, R. *et al.*, *J. Clin. Invest.* 100:503, 1997; Threlkeld, S. C. *et al.*, *J. Immunol.* 159:1648, 1997; Diepolder, H. M. *et al.*, *J. Virol.* 71:6011, 1997; Tsang *et al.*, *J. Natl. Cancer Inst.* 87:982-990, 1995; Disis *et al.*, *J. Immunol.* 156:3151-3158, 1996). In applying this strategy, recall responses are detected by culturing PBL from patients with cancer who have generated an immune response "naturally", or from patients who were vaccinated with tumor antigen vaccines. PBL from subjects are cultured *in vitro* for 1-2 weeks in the presence of test peptide plus antigen presenting cells (APC) to allow activation of "memory" T cells, as compared to "naive" T cells. At the end of the culture period, T cell activity is detected using assays for T cell activity including ^{51}Cr release involving peptide-sensitized targets, T cell proliferation, or lymphokine release.

The following describes the peptide epitopes and corresponding nucleic acids of the invention.

IV.C. Binding Affinity of Peptide Epitopes for HLA Molecules

As indicated herein, the large degree of HLA polymorphism is an important factor to be taken into account with the epitope-based approach to vaccine development. To address this factor, epitope selection encompassing identification of peptides capable of binding at high or intermediate affinity to multiple HLA molecules is preferably utilized, most preferably these epitopes bind at high or intermediate affinity to two or more allele-specific HLA molecules.

CTL-inducing peptides of interest for vaccine compositions preferably include those that have an IC_{50} or binding affinity value for class I HLA molecules of 500 nM or better (*i.e.*, the value is ≤ 500 nM). HTL-inducing peptides preferably include those that have an IC_{50} or binding affinity value for class II HLA molecules of 1000 nM or better, (*i.e.*, the value is ≤ 1,000 nM). For example, peptide binding is assessed by testing the capacity of a candidate peptide to bind to a purified HLA molecule *in vitro*. Peptides exhibiting high or intermediate affinity are then considered for further analysis. Selected peptides are tested on other members of the supertype family. In preferred embodiments, peptides that exhibit cross-reactive binding are then used in cellular screening analyses or vaccines.

As disclosed herein, higher HLA binding affinity is correlated with greater immunogenicity. Greater immunogenicity can be manifested in several different ways. Immunogenicity corresponds to whether an immune response is elicited at all, and to the vigor of any particular response, as well as to the extent of a population in which a response is elicited. For example, a peptide might elicit an immune response in a diverse array of the population, yet in no instance produce a vigorous response. Moreover, higher binding affinity peptides lead to more vigorous immunogenic responses. As a result, less peptide is required to elicit a similar biological effect if a high or intermediate affinity binding peptide is used. Thus, in preferred embodiments of the invention, high or intermediate affinity binding epitopes are particularly useful.

The relationship between binding affinity for HLA class I molecules and immunogenicity of discrete peptide epitopes on bound antigens has been determined for the first time in the art by the present inventors. The correlation between binding affinity and immunogenicity was analyzed in two different experimental approaches (*see, e.g.*, Sette, *et al., J. Immunol.* 153:5586-5592, 1994). In the first approach, the immunogenicity of potential epitopes ranging in HLA binding affinity over a 10,000-fold

range was analyzed in HLA-A*0201 transgenic mice. In the second approach, the antigenicity of approximately 100 different hepatitis B virus (HBV)-derived potential epitopes, all carrying A*0201 binding motifs, was assessed by using PBL from acute hepatitis patients. Pursuant to these approaches, it was determined that an affinity threshold value of approximately 500 nM (preferably 50 nM or less) determines the capacity of a peptide epitope to elicit a CTL response. These data are true for class I binding affinity measurements for naturally processed peptides and for synthesized T cell epitopes. These data also indicate the important role of determinant selection in the shaping of T cell responses (*see, e.g.,* Schaeffer *et al., Proc. Natl. Acad. Sci. USA* 86:4649-4653, 1989).

An affinity threshold associated with immunogenicity in the context of HLA class II DR molecules has also been delineated (*see, e.g.*, Southwood *et al. J. Immunology* 160:3363-3373,1998, and co-pending U.S.S.N. 09/009,953 filed 1/21/98). In order to define a biologically significant threshold of DR binding affinity, a database of the binding affinities of 32 DR-restricted epitopes for their restricting element (*i.e.*, the HLA molecule that binds the motif) was compiled. In approximately half of the cases (15 of 32 epitopes), DR restriction was associated with high binding affinities, *i.e.* binding affinity values of 100 nM or less. In the other half of the cases (16 of 32), DR restriction was associated with intermediate affinity (binding affinity values in the 100-1000 nM range). In only one of 32 cases was DR restriction associated with an IC_{50} of 1000 nM or greater. Thus, 1000 nM can be defined as an affinity threshold associated with immunogenicity in the context of DR molecules.

In the case of tumor-associated antigens, many CTL peptide epitopes that have been shown to induce CTL that lyse peptide-pulsed target cells and tumor cell targets endogenously expressing the epitope exhibit binding affinity or IC_{50} values of 200 nM or less. In a study that evaluated the association of binding affinity and immunogenicity of such TAA epitopes, 100% (10/10) of the high binders, *i.e.*, peptide epitopes binding at an affinity of 50 nM or less, were immunogenic and 80% (8/10) of them elicited CTLs that specifically recognized tumor cells. In the 51 to 200 nM range, very similar figures were obtained. CTL inductions positive for peptide and tumor cells were noted for 86% (6/7) and 71% (5/7) of the peptides, respectively. In the 201-500 nM range, most peptides (4/5 wildtype) were positive for induction of CTL recognizing wildtype peptide, but tumor recognition was not detected.

The binding affinity of peptides for HLA molecules can be determined as described in Example 1, below.

IV.D. Peptide Epitope Binding Motifs and Supermotifs

Through the study of single amino acid substituted antigen analogs and the sequencing of endogenously bound, naturally processed peptides, critical residues required for allele-specific binding to HLA molecules have been identified. The presence of these residues correlates with binding affinity for HLA molecules. The identification of motifs and/or supermotifs that correlate with high and intermediate affinity binding is an important issue with respect to the identification of immunogenic peptide epitopes for the inclusion in a vaccine. Kast *et al.* (*J. Immunol.* 152:3904-3912, 1994) have shown that motif-bearing peptides account for 90% of the epitopes that bind to allele-specific HLA class I molecules. In this study all possible peptides of 9 amino acids in length and overlapping by eight amino acids (240 peptides), which cover the entire sequence of the E6 and E7 proteins of human papillomavirus type 16, were evaluated for binding to five allele-specific HLA molecules that are expressed at high frequency among different ethnic groups. This unbiased set of peptides allowed an evaluation of the predictive value of HLA class I motifs. From the set of 240 peptides, 22 peptides were identified that bound to an allele-specific HLA molecule with high or intermediate affinity. Of these 22 peptides, 20 (*i.e.* 91%) were motif-bearing. Thus, this study demonstrates the value of motifs for the identification of peptide epitopes for inclusion in a vaccine: application of motif-based identification techniques will identify about 90% of the potential epitopes in a target antigen protein sequence.

Such peptide epitopes are identified in the Tables described below.

Peptides of the present invention may also comprise epitopes that bind to MHC class II DR molecules. A greater degree of heterogeneity in both size and binding frame position of the motif, relative to the N and C termini of the peptide, exists for class II peptide ligands. This increased heterogeneity of HLA class II peptide ligands is due to the structure of the binding groove of the HLA class II molecule which, unlike its class I counterpart, is open at both ends. Crystallographic analysis of HLA class II DRB*0101-peptide complexes showed that the major energy of binding is contributed by peptide residues complexed with complementary pockets on the DRB*0101 molecules. An important anchor residue engages the deepest hydrophobic pocket (*see, e.g.*, Madden, D.R. *Ann. Rev. Immunol.* 13:587, 1995) and is referred to as position 1 (P1). P1 may

represent the N-terminal residue of a class II binding peptide epitope, but more typically is flanked towards the N-terminus by one or more residues. Other studies have also pointed to an important role for the peptide residue in the 6^{th} position towards the C-terminus, relative to P1, for binding to various DR molecules.

In the past few years evidence has accumulated to demonstrate that a large fraction of HLA class I and class II molecules can be classified into a relatively few supertypes, each characterized by largely overlapping peptide binding repertoires, and consensus structures of the main peptide binding pockets. Thus, peptides of the present invention are identified by any one of several HLA-specific amino acid motifs (*see, e.g.*, Tables I-III), or if the presence of the motif corresponds to the ability to bind several allele-specific HLA molecules, a supermotif. The HLA molecules that bind to peptides that possess a particular amino acid supermotif are collectively referred to as an HLA "supertype."

The peptide motifs and supermotifs described below, and summarized in Tables I-III, provide guidance for the identification and use of peptide epitopes in accordance with the invention.

Examples of peptide epitopes bearing a respective supermotif or motif are included in Tables as designated in the description of each motif or supermotif below. The Tables include a binding affinity ratio listing for some of the peptide epitopes. The ratio may be converted to IC_{50} by using the following formula: IC_{50} of the standard peptide/ratio = IC_{50} of the test peptide (*i.e.*, the peptide epitope). The IC_{50} values of standard peptides used to determine binding affinities for Class I peptides are shown in Table IV. The IC_{50} values of standard peptides used to determine binding affinities for Class II peptides are shown in Table V. The peptides used as standards for the binding assays described herein are examples of standards; alternative standard peptides can also be used when performing binding studies.

To obtain the peptide epitope sequences listed in each Table, protein sequence data for CEA were evaluated for the presence of the designated supermotif or motif. The "pos" (position) column in the Tables designates the amino acid position in the CEA protein that corresponds to the first amino acid residue of the putative epitope. The "number of amino acids" indicates the number of residues in the epitope sequence.

HLA Class I Motifs Indicative of CTL Inducing Peptide Epitopes:

The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs delineated below are summarized in Table I. The HLA class I motifs set out in Table I(a) are those most particularly relevant to the invention claimed here. Primary and secondary anchor positions are summarized in Table II. Allele-specific HLA molecules that comprise HLA class I supertype families are listed in Table VI. In some cases, peptide epitopes may be listed in both a motif and a supermotif Table. The relationship of a particular motif and respective supermotif is indicated in the description of the individual motifs.

IV.D.1. HLA-A1 supermotif

The HLA-A1 supermotif is characterized by the presence in peptide ligands of a small (T or S) or hydrophobic (L, I, V, or M) primary anchor residue in position 2, and an aromatic (Y, F, or W) primary anchor residue at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind to the A1 supermotif (*i.e.*, the HLA-A1 supertype) is comprised of at least: A*0101, A*2601, A*2602, A*2501, and A*3201 (*see, e.g.*, DiBrino, M. *et al.*, *J. Immunol.* 151:5930, 1993; DiBrino, M. *et al.*, *J. Immunol.* 152:620, 1994; Kondo, A. *et al.*, *Immunogenetics* 45:249, 1997). Other allele-specific HLA molecules predicted to be members of the A1 superfamily are shown in Table VI. Peptides binding to each of the individual HLA proteins can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the A1 supermotif are set forth on the attached Table VII.

IV.D.2. HLA-A2 supermotif

Primary anchor specificities for allele-specific HLA-A2.1 molecules (*see, e.g.*, Falk *et al.*, *Nature* 351:290-296, 1991; Hunt *et al.*, *Science* 255:1261-1263, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992; Ruppert *et al.*, *Cell* 74:929-937, 1993) and cross-reactive binding among HLA-A2 and -A28 molecules have been described. (*See, e.g.*, Fruci *et al.*, *Human Immunol.* 38:187-192, 1993; Tanigaki *et al.*, *Human Immunol.* 39:155-162, 1994; Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994 for reviews of relevant data.) These primary anchor residues define the HLA-A2 supermotif; which presence in peptide ligands corresponds

to the ability to bind several different HLA-A2 and -A28 molecules. The HLA-A2 supermotif comprises peptide ligands with L, I, V, M, A, T, or Q as a primary anchor residue at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope.

The corresponding family of HLA molecules (*i.e.*, the HLA-A2 supertype that binds these peptides) is comprised of at least: A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, and A*6901. Other allele-specific HLA molecules predicted to be members of the A2 superfamily are shown in Table VI. As explained in detail below, binding to each of the individual allele-specific HLA molecules can be modulated by substitutions at the primary anchor and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise an A2 supermotif are set forth on the attached Table VIII. The motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

IV.D.3. HLA-A3 supermotif

The HLA-A3 supermotif is characterized by the presence in peptide ligands of A, L, I, V, M, S, or, T as a primary anchor at position 2, and a positively charged residue, R or K, at the C-terminal position of the epitope, *e.g.*, in position 9 of 9-mers (*see, e.g.*, Sidney *et al.*, *Hum. Immunol.* 45:79, 1996). Exemplary members of the corresponding family of HLA molecules (the HLA-A3 supertype) that bind the A3 supermotif include at least: A*0301, A*1101, A*3101, A*3301, and A*6801. Other allele-specific HLA molecules predicted to be members of the A3 supertype are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions of amino acids at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the A3 supermotif are set forth on the attached Table IX.

IV.D.4. HLA-A24 supermotif

The HLA-A24 supermotif is characterized by the presence in peptide ligands of an aromatic (F, W, or Y) or hydrophobic aliphatic (L, I, V, M, or T) residue as a primary anchor in position 2, and Y, F, W, L, I, or M as primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sette and Sidney, *Immunogenetics*, in press, 1999). The corresponding family of HLA molecules that bind to the A24 supermotif (*i.e.*, the A24 supertype) includes at least: A*2402, A*3001, and A*2301. Other allele-specific HLA molecules predicted to be members of the A24 supertype are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the A24 supermotif are set forth on the attached Table X.

IV.D.5. HLA-B7 supermotif

The HLA-B7 supermotif is characterized by peptides bearing proline in position 2 as a primary anchor, and a hydrophobic or aliphatic amino acid (L, I, V, M, A, F, W, or Y) as the primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind the B7 supermotif (*i.e.*, the HLA-B7 supertype) is comprised of at least twenty six HLA-B proteins comprising at least: B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, and B*7801 (*see, e.g.*, Sidney, *et al., J. Immunol.* 154:247, 1995; Barber, *et al., Curr. Biol.* 5:179, 1995; Hill, *et al., Nature* 360:434, 1992; Rammensee, *et al., Immunogenetics* 41:178, 1995 for reviews of relevant data). Other allele-specific HLA molecules predicted to be members of the B7 supertype are shown in Table VI. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the B7 supermotif are set forth on the attached Table XI.

IV.D.6. HLA-B27 supermotif

The HLA-B27 supermotif is characterized by the presence in peptide ligands of a positively charged (R, H, or K) residue as a primary anchor at position 2, and a hydrophobic (F, Y, L, W, M, I, A, or V) residue as a primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sidney and Sette, *Immunogenetics*, in press, 1999). Exemplary members of the corresponding family of HLA molecules that bind to the B27 supermotif (*i.e.*, the B27 supertype) include at least B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, and B*7301. Other allele-specific HLA molecules predicted to be members of the B27 supertype are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the B27 supermotif are set forth on the attached Table XII.

IV.D.7. HLA-B44 supermotif

The HLA-B44 supermotif is characterized by the presence in peptide ligands of negatively charged (D or E) residues as a primary anchor in position 2, and hydrophobic residues (F, W, Y, L, I, M, V, or A) as a primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sidney et al., *Immunol. Today* 17:261, 1996). Exemplary members of the corresponding family of HLA molecules that bind to the B44 supermotif (*i.e.*, the B44 supertype) include at least: B*1801, B*1802, B*3701, B*4001, B*4002, B*4006, B*4402, B*4403, and B*4404. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the supermotif.

IV.D.8. HLA-B58 supermotif

The HLA-B58 supermotif is characterized by the presence in peptide ligands of a small aliphatic residue (A, S, or T) as a primary anchor residue at position 2, and an aromatic or hydrophobic residue (F, W, Y, L, I, V, M, or A) as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, Sidney and Sette, *Immunogenetics*, in press, 1999 for reviews of relevant data). Exemplary members of the corresponding family of HLA molecules that bind to the B58 supermotif (*i.e.*, the B58 supertype) include at least: B*1516, B*1517, B*5701, B*5702, and B*5801. Other allele-specific

HLA molecules predicted to be members of the B58 supertype are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the B58 supermotif are set forth on the attached Table XIII.

IV.D.9. HLA-B62 supermotif

The HLA-B62 supermotif is characterized by the presence in peptide ligands of the polar aliphatic residue Q or a hydrophobic aliphatic residue (L, V, M, I, or P) as a primary anchor in position 2, and a hydrophobic residue (F, W, Y, M, I, V, L, or A) as a primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sidney and Sette, *Immunogenetics*, in press, 1999). Exemplary members of the corresponding family of HLA molecules that bind to the B62 supermotif (*i.e.*, the B62 supertype) include at least: B*1501, B*1502, B*1513, and B5201. Other allele-specific HLA molecules predicted to be members of the B62 supertype are shown in Table VI. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Representative peptide epitopes that comprise the B62 supermotif are set forth on the attached Table XIV.

IV.D.10. HLA-A1 motif

The HLA-A1 motif is characterized by the presence in peptide ligands of T, S, or M as a primary anchor residue at position 2 and the presence of Y as a primary anchor residue at the C-terminal position of the epitope. An alternative allele-specific A1 motif is characterized by a primary anchor residue at position 3 rather than position 2. This motif is characterized by the presence of D, E, A, or S as a primary anchor residue in position 3, and a Y as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, DiBrino *et al.*, J. Immunol., 152:620, 1994; Kondo *et al.*, *Immunogenetics* 45:249, 1997; and Kubo *et al.*, *J. Immunol.* 152:3913, 1994 for reviews of relevant data). Peptide binding to HLA-A1 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise either A1 motif are set forth on the attached Table XV. Those epitopes comprising T, S, or M at position 2 and Y at the C-terminal position are also included in the listing of HLA-A1 supermotif-bearing peptide epitopes listed in Table VII, as these residues are a subset of the A1 supermotif primary anchors.

IV.D.11. HLA-A*0201 motif

An HLA-A2*0201 motif was determined to be characterized by the presence in peptide ligands of L or M as a primary anchor residue in position 2, and L or V as a primary anchor residue at the C-terminal position of a 9-residue peptide (*see, e.g.*, Falk *et al.*, *Nature* 351:290-296, 1991) and was further found to comprise an I at position 2 and I or A at the C-terminal position of a nine amino acid peptide (*see, e.g.*, Hunt *et al.*, Science 255:1261-1263, March 6, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992). The A*0201 allele-specific motif has also been defined by the present inventors to additionally comprise V, A, T, or Q as a primary anchor residue at position 2, and M or T as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). Thus, the HLA-A*0201 motif comprises peptide ligands with L, I, V, M, A, T, or Q as primary anchor residues at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope. The preferred and tolerated residues that characterize the primary anchor positions of the HLA-A*0201 motif are identical to the residues describing the A2 supermotif. (For reviews of relevant data, *see, e.g.*, del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Ruppert *et al.*, *Cell* 74:929-937, 1993; Sidney *et al.*, *Immunol. Today* 17:261-266, 1996; Sette and Sidney, *Curr. Opin. in Immunol.* 10:478-482, 1998). Secondary anchor residues that characterize the A*0201 motif have additionally been defined (*see, e.g.*, Ruppert *et al.*, *Cell* 74:929-937, 1993). These are shown in Table II. Peptide binding to HLA-A*0201 molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise an A*0201 motif are set forth on the attached Table VIII. The A*0201 motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

IV.D.12. HLA-A3 motif

The HLA-A3 motif is characterized by the presence in peptide ligands of L, M, V, I, S, A, T, F, C, G, or D as a primary anchor residue at position 2, and the presence of K, sY, R, H, F, or A as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, DiBrino *et al.*, *Proc. Natl. Acad. Sci USA* 90:1508, 1993; and Kubo *et al.*, *J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise the A3 motif are set forth on the attached Table XVI. Those peptide epitopes that also comprise the A3 supermotif are also listed in Table IX. The A3 supermotif primary anchor residues comprise a subset of the A3- and A11-allele specific motif primary anchor residues.

IV.D.13. HLA-A11 motif

The HLA-A11 motif is characterized by the presence in peptide ligands of V, T, M, L, I, S, A, G, N, C, D, or F as a primary anchor residue in position 2, and K, R, Y, or H as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, Zhang *et al.*, *Proc. Natl. Acad. Sci USA* 90:2217-2221, 1993; and Kubo *et al.*, *J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A11 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise the A11 motif are set forth on the attached Table XVII; peptide epitopes comprising the A3 allele-specific motif are also present in this Table because of the extensive overlap between the A3 and A11 motif primary anchor specificities. Further, those peptide epitopes that comprise the A3 supermotif are also listed in Table IX.

IV.D.14. HLA-A24 motif

The HLA-A24 motif is characterized by the presence in peptide ligands of Y, F, W, or M as a primary anchor residue in position 2, and F, L, I, or W as a primary anchor residue at the C-terminal position of the epitope (*see, e.g., Kondo et al., J. Immunol.* 155:4307-4312, 1995; and Kubo *et al.*, *J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A24 molecules can be modulated by substitutions at primary and/or

secondary anchor positions; preferably choosing respective residues specified for the motif.

Representative peptide epitopes that comprise the A24 motif are set forth on the attached Table XVIII. These epitopes are also listed in Table X, which sets forth HLA-A24-supermotif-bearing peptide epitopes, as the primary anchor residues characterizing the A24 allele-specific motif comprise a subset of the A24 supermotif primary anchor residues.

Motifs Indicative of Class II HTL Inducing Peptide Epitopes

The primary and secondary anchor residues of the HLA class II peptide epitope supermotifs and motifs delineated below are summarized in Table III.

IV.D.15. HLA DR-1-4-7 supermotif

Motifs have also been identified for peptides that bind to three common HLA class II allele-specific HLA molecules: HLA DRB1*0401, DRB1*0101, and DRB1*0701 (*see, e.g.*, the review by Southwood *et al. J. Immunology* 160:3363-3373,1998). Collectively, the common residues from these motifs delineate the HLA DR-1-4-7 supermotif. Peptides that bind to these DR molecules carry a supermotif characterized by a large aromatic or hydrophobic residue (Y, F, W, L, I, V, or M) as a primary anchor residue in position 1, and a small, non-charged residue (S, T, C, A, P, V, I, L, or M) as a primary anchor residue in position 6 of a 9-mer core region. Allele-specific secondary effects and secondary anchors for each of these HLA types have also been identified (Southwood *et al., supra*). These are set forth in Table III. Peptide binding to HLA-DRB1*0401, DRB1*0101, and/or DRB1*0701 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

Potential epitope 9-mer core regions comprising the DR-1-4-7 supermotif, wherein position 1 of the supermotif is at position 1 of the nine-residue core, are set forth in Table XIX. Respective exemplary peptide epitopes of 15 amino acid residues in length, each of which comprise the nine residue core, are also shown in the Table along with cross-reactive binding data for the exemplary 15-residue supermotif-bearing peptides.

IV.D.16. HLA DR3 motifs

Two alternative motifs (*i.e.*, submotifs) characterize peptide epitopes that bind to HLA-DR3 molecules (*see, e.g.*, Geluk *et al.*, *J. Immunol.* 152:5742, 1994). In the first motif (submotif DR3a) a large, hydrophobic residue (L, I, V, M, F, or Y) is present in anchor position 1 of a 9-mer core, and D is present as an anchor at position 4, towards the carboxyl terminus of the epitope. As in other class II motifs, core position 1 may or may not occupy the peptide N-terminal position.

The alternative DR3 submotif provides for lack of the large, hydrophobic residue at anchor position 1, and/or lack of the negatively charged or amide-like anchor residue at position 4, by the presence of a positive charge at position 6 towards the carboxyl terminus of the epitope. Thus, for the alternative allele-specific DR3 motif (submotif DR3b): L, I, V, M, F, Y, A, or Y is present at anchor position 1; D, N, Q, E, S, or T is present at anchor position 4; and K, R, or H is present at anchor position 6. Peptide binding to HLA-DR3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

Potential peptide epitope 9-mer core regions corresponding to a nine residue sequence comprising the DR3a submotif (wherein position 1 of the motif is at position 1 of the nine residue core) are set forth in Table XXa. Respective exemplary peptide epitopes of 15 amino acid residues in length, each of which comprise the nine residue core, are also shown in Table XXa along with binding data for exemplary DR3 submotif a-bearing peptides.

Potential peptide epitope 9-mer core regions comprising the DR3b submotif and respective exemplary 15-mer peptides comprising the DR3 submotif-b epitope are set forth in Table XXb along with binding data of exemplary DR3 submotif b-bearing peptides.

Each of the HLA class I or class II peptide epitopes set out in the Tables herein are deemed singly to be an inventive aspect of this application. Further, it is also an inventive aspect of this application that each peptide epitope may be used in combination with any other peptide epitope.

IV.E. Enhancing Population Coverage of the Vaccine

Vaccines that have broad population coverage are preferred because they are more commercially viable and generally applicable to the most people. Broad population coverage can be obtained using the peptides of the invention (and nucleic acid

compositions that encode such peptides) through selecting peptide epitopes that bind to HLA alleles which, when considered in total, are present in most of the population. Table XXI lists the overall frequencies of the HLA class I supertypes in various ethnicities (Table XXIa) and the combined population coverage achieved by the A2-, A3-, and B7-supertypes (Table XXIb). The A2-, A3-, and B7 supertypes are each present on the average of over 40% in each of these five major ethnic groups. Coverage in excess of 80% is achieved with a combination of these supermotifs. These results suggest that effective and non-ethnically biased population coverage is achieved upon use of a limited number of cross-reactive peptides. Although the population coverage reached with these three main peptide specificities is high, coverage can be expanded to reach 95% population coverage and above, and more easily achieve truly multispecific responses upon use of additional supermotif or allele-specific motif bearing peptides.

The B44-, A1-, and A24-supertypes are each present, on average, in a range from 25% to 40% in these major ethnic populations (Table XXIa). While less prevalent overall, the B27-, B58-, and B62 supertypes are each present with a frequency >25% in at least one major ethnic group (Table XXIa). Table XXIb summarizes the estimated prevalence of combinations of HLA supertypes that have been identified in five major ethnic groups. The incremental coverage obtained by the inclusion of A1,- A24-, and B44-supertypes to the A2, A3, and B7 coverage and coverage obtained with all of the supertypes described herein, is shown.

The data presented herein, together with the previous definition of the A2-, A3-, and B7-supertypes, indicates that all antigens, with the possible exception of A29, B8, and B46, can be classified into a total of nine HLA supertypes. By including epitopes from the six most frequent supertypes, an average population coverage of 99% is obtained for five major ethnic groups.

IV.F. Immune Response-Stimulating Peptide Analogs

In general, CTL and HTL responses are not directed against all possible epitopes. Rather, they are restricted to a few "immunodominant" determinants (Zinkernagel, *et al., Adv. Immunol.* 27:5159, 1979; Bennink, *et al., J. Exp. Med.* 168:19351939, 1988; Rawle, *et al., J. Immunol.* 146:3977-3984, 1991). It has been recognized that immunodominance (Benacerraf, *et al., Science* 175:273-279, 1972) could be explained by either the ability of a given epitope to selectively bind a particular HLA protein (determinant selection theory) (Vitiello, *et al., J. Immunol.* 131:1635, 1983); Rosenthal, *et al., Nature* 267:156-

158, 1977), or to be selectively recognized by the existing TCR (T cell receptor) specificities (repertoire theory) (Klein, J., IMMUNOLOGY, THE SCIENCE OF SELF/NONSELF DISCRIMINATION, John Wiley & Sons, New York, pp. 270-310, 1982). It has been demonstrated that additional factors, mostly linked to processing events, can also play a key role in dictating, beyond strict immunogenicity, which of the many potential determinants will be presented as immunodominant (Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993).

Because tissue specific and developmental TAAs are expressed on normal tissue at least at some point in time or location within the body, it may be expected that T cells to them, particularly dominant epitopes, are eliminated during immunological surveillance and that tolerance is induced. However, CTL responses to tumor epitopes in both normal donors and cancer patient has been detected, which may indicate that tolerance is incomplete (*see, e.g.*, Kawashima *et al.*, Hum. Immunol. 59:1, 1998; Tsang, *J. Natl. Cancer Inst.* 87:82-90, 1995; Rongcun *et al.*, *J. Immunol.* 163:1037, 1999). Thus, immune tolerance does not completely eliminate or inactivate CTL precursors capable of recognizing high affinity HLA class I binding peptides.

An additional strategy to overcome tolerance is to use analog peptides. Without intending to be bound by theory, it is believed that because T cells to dominant epitopes may have been clonally deleted, selecting subdominant epitopes may allow existing T cells to be recruited, which will then lead to a therapeutic or prophylactic response. However, the binding of HLA molecules to subdominant epitopes is often less vigorous than to dominant ones. Accordingly, there is a need to be able to modulate the binding affinity of particular immunogenic epitopes for one or more HLA molecules, and thereby to modulate the immune response elicited by the peptide, for example to prepare analog peptides which elicit a more vigorous response.

Although peptides with suitable cross-reactivity among all alleles of a superfamily are identified by the screening procedures described above, cross-reactivity is not always as complete as possible, and in certain cases procedures to increase cross-reactivity of peptides can be useful; moreover, such procedures can also be used to modify other properties of the peptides such as binding affinity or peptide stability. Having established the general rules that govern cross-reactivity of peptides for HLA alleles within a given motif or supermotif, modification (*i.e.*, analoging) of the structure of peptides of particular interest in order to achieve broader (or otherwise modified) HLA binding capacity can be performed. More specifically, peptides which exhibit the broadest cross-

reactivity patterns, can be produced in accordance with the teachings herein. The present concepts related to analog generation are set forth in greater detail in co-pending U.S.S.N. 09/226,775 filed 1/6/99.

In brief, the strategy employed utilizes the motifs or supermotifs which correlate with binding to certain HLA molecules. The motifs or supermotifs are defined by having primary anchors, and in many cases secondary anchors. Analog peptides can be created by substituting amino acid residues at primary anchor, secondary anchor, or at primary and secondary anchor positions. Generally, analogs are made for peptides that already bear a motif or supermotif. Preferred secondary anchor residues of supermotifs and motifs that have been defined for HLA class I and class II binding peptides are shown in Tables II and III, respectively.

For a number of the motifs or supermotifs in accordance with the invention, residues are defined which are deleterious to binding to allele-specific HLA molecules or members of HLA supertypes that bind the respective motif or supermotif (Tables II and III). Accordingly, removal of such residues that are detrimental to binding can be performed in accordance with the present invention. For example, in the case of the A3 supertype, when all peptides that have such deleterious residues are removed from the population of peptides used in the analysis, the incidence of cross-reactivity increased from 22% to 37% (*see, e.g.*, Sidney, J. *et al.*, *Hu. Immunol.* 45:79, 1996). Thus, one strategy to improve the cross-reactivity of peptides within a given supermotif is simply to delete one or more of the deleterious residues present within a peptide and substitute a small "neutral" residue such as Ala (that may not influence T cell recognition of the peptide). An enhanced likelihood of cross-reactivity is expected if, together with elimination of detrimental residues within a peptide, "preferred" residues associated with high affinity binding to an allele-specific HLA molecule or to multiple HLA molecules within a superfamily are inserted.

To ensure that an analog peptide, when used as a vaccine, actually elicits a CTL response to the native epitope *in vivo* (or, in the case of class II epitopes, elicits helper T cells that cross-react with the wild type peptides), the analog peptide may be used to immunize T cells *in vitro* from individuals of the appropriate HLA allele. Thereafter, the immunized cells' capacity to induce lysis of wild type peptide sensitized target cells is evaluated. It will be desirable to use as antigen presenting cells, cells that have been either infected, or transfected with the appropriate genes, or, in the case of class II

epitopes only, cells that have been pulsed with whole protein antigens, to establish whether endogenously produced antigen is also recognized by the relevant T cells.

Another embodiment of the invention is to create analogs of weak binding peptides, to thereby ensure adequate numbers of cross-reactive cellular binders. Class I binding peptides exhibiting binding affinities of 500-5000 nM, and carrying an acceptable but suboptimal primary anchor residue at one or both positions can be "fixed" by substituting preferred anchor residues in accordance with the respective supertype. The analog peptides can then be tested for crossbinding activity.

Another embodiment for generating effective peptide analogs involves the substitution of residues that have an adverse impact on peptide stability or solubility in, *e.g.*, a liquid environment. This substitution may occur at any position of the peptide epitope. For example, a cysteine can be substituted out in favor of α-amino butyric acid ("B" in the single letter abbreviations for peptide sequences listed herein). Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substituting α-amino butyric acid for cysteine not only alleviates this problem, but actually improves binding and crossbinding capability in certain instances (*see, e.g.*, the review by Sette *et al.*, In: Persistent Viral Infections, Eds. R. Ahmed and I. Chen, John Wiley & Sons, England, 1999).

Representative analog peptides are set forth in Table XXII. The Table indicates the length and sequence of the analog peptide as well as the motif or supermotif, if appropriate. The information in the "Fixed Nomenclature" column indicates the residues substituted at the indicated position numbers for the respective analog.

IV.G. Computer Screening of Protein Sequences from Disease-Related Antigens for Supermotif- or Motif-Bearing Peptides

In order to identify supermotif- or motif-bearing epitopes in a target antigen, a native protein sequence, *e.g.*, a tumor-associated antigen, or sequences from an infectious organism, or a donor tissue for transplantation, is screened using a means for computing, such as an intellectual calculation or a computer, to determine the presence of a supermotif or motif within the sequence. The information obtained from the analysis of native peptide can be used directly to evaluate the status of the native peptide or may be utilized subsequently to generate the peptide epitope.

Computer programs that allow the rapid screening of protein sequences for the occurrence of the subject supermotifs or motifs are encompassed by the present invention; as are programs that permit the generation of analog peptides. These programs are implemented to analyze any identified amino acid sequence or operate on an unknown sequence and simultaneously determine the sequence and identify motif-bearing epitopes thereof; analogs can be simultaneously determined as well. Generally, the identified sequences will be from a pathogenic organism or a tumor-associated peptide. For example, the target TAA molecules include, without limitation, CEA, MAGE, p53 and her2/neu.

It is important that the selection criteria utilized for prediction of peptide binding are as accurate as possible, to correlate most efficiently with actual binding. Prediction of peptides that bind, for example, to HLA-A*0201, on the basis of the presence of the appropriate primary anchors, is positive at about a 30% rate (*see, e.g.*, Ruppert, J. *et al. Cell* 74:929, 1993). However, by extensively analyzing peptide-HLA binding data disclosed herein, data in related patent applications, and data in the art, the present inventors have developed a number of allele-specific polynomial algorithms that dramatically increase the predictive value over identification on the basis of the presence of primary anchor residues alone. These algorithms take into account not only the presence or absence of primary anchors, but also consider the positive or deleterious presence of secondary anchor residues (to account for the impact of different amino acids at different positions). The algorithms are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA interactions can be approximated as a linear polynomial function of the type:

$$\Delta G = a_{1i} \times a_{2i} \times a_{3i} \ldots \times a_{ni}$$

where a_{ji} is a coefficient that represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. An important assumption of this method is that the effects at each position are essentially independent of each other. This assumption is justified by studies that demonstrated that peptides are bound to HLA molecules and recognized by T cells in essentially an extended conformation. Derivation of specific algorithm coefficients has been described, for example, in Gulukota, K. *et al.*, *J. Mol. Biol.* 267:1258, 1997.

Additional methods to identify preferred peptide sequences, which also make use of specific motifs, include the use of neural networks and molecular modeling programs

(*see, e.g.*, Milik *et al.*, *Nature Biotechnology* 16:753, 1998; Altuvia *et al.*, *Hum. Immunol.* 58:1, 1997; Altuvia *et al*, *J. Mol. Biol.* 249:244, 1995; Buus, S. *Curr. Opin. Immunol.* 11:209-213, 1999; Brusic, V. *et al.*, *Bioinformatics* 14:121-130, 1998; Parker *et al.*, *J. Immunol.* 152:163, 1993; Meister *et al.*, *Vaccine* 13:581, 1995; Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994; Sturniolo *et al.*, *Nature Biotechnol.* 17:555 1999).

For example, it has been shown that in sets of A*0201 motif-bearing peptides containing at least one preferred secondary anchor residue while avoiding the presence of any deleterious secondary anchor residues, 69% of the peptides will bind A*0201 with an IC_{50} less than 500 nM (Ruppert, J. *et al. Cell* 74:929, 1993). These algorithms are also flexible in that cut-off scores may be adjusted to select sets of peptides with greater or lower predicted binding properties, as desired.

In utilizing computer screening to identify peptide epitopes, a protein sequence or translated sequence may be analyzed using software developed to search for motifs, for example the "FINDPATTERNS' program (Devereux, *et al. Nucl. Acids Res.* 12:387-395, 1984) or MotifSearch 1.4 software program (D. Brown, San Diego, CA) to identify potential peptide sequences containing appropriate HLA binding motifs. The identified peptides can be scored using customized polynomial algorithms to predict their capacity to bind specific HLA class I or class II alleles. As appreciated by one of ordinary skill in the art, a large array of computer programming software and hardware options are available in the relevant art which can be employed to implement the motifs of the invention in order to evaluate (*e.g.*, without limitation, to identify epitopes, identify epitope concentration per peptide length, or to generate analogs) known or unknown peptide sequences.

In accordance with the procedures described above, CEA peptide epitopes and analogs thereof that are able to bind HLA supertype groups or allele-specific HLA molecules have been identified (Tables VII-XX; Table XXII).

IV.H. Preparation of Peptide Epitopes

Peptides in accordance with the invention can be prepared synthetically, by recombinant DNA technology or chemical synthesis, or from natural sources such as native tumors or pathogenic organisms. Peptide epitopes may be synthesized individually or as polyepitopic peptides. Although the peptide will preferably be substantially free of other naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides may be synthetically conjugated to native fragments or particles.

The peptides in accordance with the invention can be a variety of lengths, and either in their neutral (uncharged) forms or in forms which are salts. The peptides in accordance with the invention are either free of modifications such as glycosylation, side chain oxidation, or phosphorylation; or they contain these modifications, subject to the condition that modifications do not destroy the biological activity of the peptides as described herein.

Desirably, the peptide epitope will be as small as possible while still maintaining substantially all of the immunologic activity of the native protein. When possible, it may be desirable to optimize HLA class I binding peptide epitopes of the invention to a length of about 8 to about 13 amino acid residues, preferably 9 to 10. HLA class II binding peptide epitopes may be optimized to a length of about 6 to about 30 amino acids in length, preferably to between about 13 and about 20 residues. Preferably, the peptide epitopes are commensurate in size with endogenously processed pathogen-derived peptides or tumor cell peptides that are bound to the relevant HLA molecules.

The identification and preparation of peptides of other lengths can also be carried out using the techniques described herein. Moreover, it is preferred to identify native peptide regions that contain a high concentration of class I and/or class II epitopes. Such a sequence is generally selected on the basis that it contains the greatest number of epitopes per amino acid length. It is to be appreciated that epitopes can be present in a frame-shifted manner, *e.g.* a 10 amino acid long peptide could contain two 9 amino acid long epitopes and one 10 amino acid long epitope; upon intracellular processing, each epitope can be exposed and bound by an HLA molecule upon administration of such a peptide. This larger, preferably multi-epitopic, peptide can be generated synthetically, recombinantly, or via cleavage from the native source.

The peptides of the invention can be prepared in a wide variety of ways. For the preferred relatively short size, the peptides can be synthesized in solution or on a solid support in accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. (*See*, for example, Stewart & Young, SOLID PHASE PEPTIDE SYNTHESIS, 2D. ED., Pierce Chemical Co., 1984). Further, individual peptide epitopes can be joined using chemical ligation to produce larger peptides that are still within the bounds of the invention.

Alternatively, recombinant DNA technology can be employed wherein a nucleotide sequence which encodes an immunogenic peptide of interest is inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated

under conditions suitable for expression. These procedures are generally known in the art, as described generally in Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989). Thus, recombinant polypeptides which comprise one or more peptide sequences of the invention can be used to present the appropriate T cell epitope.

The nucleotide coding sequence for peptide epitopes of the preferred lengths contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al., J. Am. Chem. Soc.* 103:3185 (1981). Peptide analogs can be made simply by substituting the appropriate and desired nucleic acid base(s) for those that encode the native peptide sequence; exemplary nucleic acid substitutions are those that encode an amino acid defined by the motifs/supermotifs herein. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly available in the art, and the vectors used to transform suitable hosts to produce the desired fusion protein. A number of such vectors and suitable host systems are now available. For expression of the fusion proteins, the coding sequence will be provided with operably linked start and stop codons, promoter and terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are transformed into suitable bacterial hosts. Of course, yeast, insect or mammalian cell hosts may also be used, employing suitable vectors and control sequences.

IV.I. Assays to Detect T-Cell Responses

Once HLA binding peptides are identified, they can be tested for the ability to elicit a T-cell response. The preparation and evaluation of motif-bearing peptides are described in PCT publications WO 94/20127 and WO 94/03205. Briefly, peptides comprising epitopes from a particular antigen are synthesized and tested for their ability to bind to the appropriate HLA proteins. These assays may involve evaluating the binding of a peptide of the invention to purified HLA class I molecules in relation to the binding of a radioiodinated reference peptide. Alternatively, cells expressing empty class I molecules (*i.e.* lacking peptide therein) may be evaluated for peptide binding by immunofluorescent staining and flow microfluorimetry. Other assays that may be used to

evaluate peptide binding include peptide-dependent class I assembly assays and/or the inhibition of CTL recognition by peptide competition. Those peptides that bind to the class I molecule, typically with an affinity of 500 nM or less, are further evaluated for their ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary *in vitro* or *in vivo* CTL responses that can give rise to CTL populations capable of reacting with selected target cells associated with a disease. Corresponding assays are used for evaluation of HLA class II binding peptides. HLA class II motif-bearing peptides that are shown to bind, typically at an affinity of 1000 nM or less, are further evaluated for the ability to stimulate HTL responses.

Conventional assays utilized to detect T cell responses include proliferation assays, lymphokine secretion assays, direct cytotoxicity assays, and limiting dilution assays. For example, antigen-presenting cells that have been incubated with a peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells. Alternatively, mutant non-human mammalian cell lines that are deficient in their ability to load class I molecules with internally processed peptides and that have been transfected with the appropriate human class I gene, may be used to test for the capacity of the peptide to induce *in vitro* primary CTL responses.

Peripheral blood mononuclear cells (PBMCs) may be used as the responder cell source of CTL precursors. The appropriate antigen-presenting cells are incubated with peptide, after which the peptide-loaded antigen-presenting cells are then incubated with the responder cell population under optimized culture conditions. Positive CTL activation can be determined by assaying the culture for the presence of CTLs that kill radio-labeled target cells, both specific peptide-pulsed targets as well as target cells expressing endogenously processed forms of the antigen from which the peptide sequence was derived.

More recently, a method has been devised which allows direct quantification of antigen-specific T cells by staining with Fluorescein-labelled HLA tetrameric complexes (Altman, J. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:10330, 1993; Altman, J. D. *et al.*, *Science* 274:94, 1996). Other relatively recent technical developments include staining for intracellular lymphokines, and interferon-γ release assays or ELISPOT assays. Tetramer staining, intracellular lymphokine staining and ELISPOT assays all appear to be at least 10-fold more sensitive than more conventional assays (Lalvani, A. *et al.*, *J. Exp.*

Med. 186:859, 1997; Dunbar, P. R. *et al.*, *Curr. Biol.* 8:413, 1998; Murali-Krishna, K. *et al.*, *Immunity* 8:177, 1998).

HTL activation may also be assessed using such techniques known to those in the art such as T cell proliferation and secretion of lymphokines, *e.g.* IL-2 (*see*, *e.g.* Alexander *et al.*, Immunity 1:751-761, 1994).

Alternatively, immunization of HLA transgenic mice can be used to determine immunogenicity of peptide epitopes. Several transgenic mouse models including mice with human A2.1, A11 (which can additionally be used to analyze HLA-A3 epitopes), and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed. Additional transgenic mouse models with other HLA alleles may be generated as necessary. Mice may be immunized with peptides emulsified in Incomplete Freund's Adjuvant and the resulting T cells tested for their capacity to recognize peptide-pulsed target cells and target cells transfected with appropriate genes. CTL responses may be analyzed using cytotoxicity assays described above. Similarly, HTL responses may be analyzed using such assays as T cell proliferation or secretion of lymphokines.

Exemplary immunogenic peptide epitopes are set out in Table XXIII.

IV.J. Use of Peptide Epitopes as Diagnostic Agents and for Evaluating Immune Responses

HLA class I and class II binding peptides as described herein can be used, in one embodiment of the invention, as reagents to evaluate an immune response. The immune response to be evaluated may be induced by using as an immunogen any agent that may result in the production of antigen-specific CTLs or HTLs that recognize and bind to the peptide epitope(s) to be employed as the reagent. The peptide reagent need not be used as the immunogen. Assay systems that may be used for such an analysis include relatively recent technical developments such as tetramers, staining for intracellular lymphokines and interferon release assays, or ELISPOT assays.

For example, a peptide of the invention may be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of antigen-specific CTLs following exposure to a tumor cell antigen or an immunogen. The HLA-tetrameric complex is used to directly visualize antigen-specific CTLs (*see*, *e.g.*, Ogg *et al.*, *Science* 279:2103-2106, 1998; and Altman *et al.*, *Science* 174:94-96, 1996) and determine the frequency of the antigen-specific CTL population in a sample of peripheral blood

mononuclear cells. A tetramer reagent using a peptide of the invention may be generated as follows: A peptide that binds to an HLA molecule is refolded in the presence of the corresponding HLA heavy chain and β_2-microglobulin to generate a trimolecular complex. The complex is biotinylated at the carboxyl terminal end of the heavy chain at a site that was previously engineered into the protein. Tetramer formation is then induced by the addition of streptavidin. By means of fluorescently labeled streptavidin, the tetramer can be used to stain antigen-specific cells. The cells may then be identified, for example, by flow cytometry. Such an analysis may be used for diagnostic or prognostic purposes.

Peptides of the invention may also be used as reagents to evaluate immune recall responses (*see, e.g.*, Bertoni *et al.*, *J. Clin. Invest.* 100:503-513, 1997 and Penna *et al.*, *J. Exp. Med.* 174:1565-1570, 1991). For example, patient PBMC samples from individuals with cancer may be analyzed for the presence of antigen-specific CTLs or HTLs using specific peptides. A blood sample containing mononuclear cells may be evaluated by cultivating the PBMCs and stimulating the cells with a peptide of the invention. After an appropriate cultivation period, the expanded cell population may be analyzed, for example, for CTL or for HTL activity.

The peptides may also be used as reagents to evaluate the efficacy of a vaccine. PBMCs obtained from a patient vaccinated with an immunogen may be analyzed using, for example, either of the methods described above. The patient is HLA typed, and peptide epitope reagents that recognize the allele-specific molecules present in that patient are selected for the analysis. The immunogenicity of the vaccine is indicated by the presence of epitope-specific CTLs and/or HTLs in the PBMC sample.

The peptides of the invention may also be used to make antibodies, using techniques well known in the art (see, *e.g.* *CURRENT PROTOCOLS IN IMMUNOLOGY*, Wiley/Greene, NY; and *Antibodies A Laboratory Manual,* Harlow and Lane, Cold Spring Harbor Laboratory Press, 1989), which may be useful as reagents to diagnose or monitor cancer. Such antibodies include those that recognize a peptide in the context of an HLA molecule, *i.e.*, antibodies that bind to a peptide-MHC complex.

IV.K. Vaccine Compositions

Vaccines that contain an immunogenically effective amount of one or more peptides as described herein are a further embodiment of the invention. Once

appropriately immunogenic epitopes have been defined, they can be sorted and delivered by various means, herein referred to as "vaccine" compositions. Such vaccine compositions can include, for example, lipopeptides (*e.g.*,Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), peptide compositions encapsulated in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991: Alonso *et al.*, *Vaccine* 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995), peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998), multiple antigen peptide systems (MAPs) (*see e.g.*, Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996), viral delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990), particles of viral or synthetic origin (*e.g.*, Kofler, N. *et al.*, *J. Immunol. Methods.* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995), adjuvants (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993), liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996), or, naked or particle absorbed cDNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993). Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) may also be used.

Furthermore, vaccines in accordance with the invention encompass compositions of one or more of the claimed peptide(s). The peptide(s) can be individually linked to its own carrier; alternatively, the peptide(s) can exist as a homopolymer or heteropolymer of active peptide units. Such a polymer has the advantage of increased immunological reaction and, where different peptide epitopes are used to make up the polymer, the additional ability to induce antibodies and/or CTLs that react with different antigenic determinants of the pathogenic organism or tumor-related peptide targeted for an immune

response. The composition may be a naturally occurring region of an antigen or may be prepared, *e.g.*, recombinantly or by chemical synthesis.

Furthermore, useful carriers that can be used with vaccines of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza, hepatitis B virus core protein, and the like. The vaccines can contain a physiologically tolerable (*i.e.*, acceptable) diluent such as water, or saline, preferably phosphate buffered saline. The vaccines also typically include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glycerylcysteinlyseryl- serine (P_3CSS).

As disclosed in greater detail herein, upon immunization with a peptide composition in accordance with the invention, via injection, aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing large amounts of CTLs and/or HTLs specific for the desired antigen. Consequently, the host becomes at least partially immune to later infection, or at least partially resistant to developing an ongoing chronic infection, or derives at least some therapeutic benefit when the antigen was tumor-associated.

In some instances it may be desirable to combine the class I peptide vaccines of the invention with vaccines which induce or facilitate neutralizing antibody responses to the target antigen of interest, particularly to viral envelope antigens. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. An alternative embodiment of such a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE™ (Epimmune, San Diego, CA) molecule (described, for example, in U.S. Patent Number 5,736,142). Furthermore, any of these embodiments can be administered as a nucleic acid mediated modality.

For therapeutic or prophylactic immunization purposes, the peptides of the invention can also be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. This approach involves the use of vaccinia virus, for example, as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into a host

bearing a tumor, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits a host CTL and/or HTL response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein.

Antigenic peptides are used to elicit a CTL and/or HTL response *ex vivo*, as well. The resulting CTL or HTL cells, can be used to treat chronic infections, or tumors in patients that do not respond to other conventional forms of therapy, or will not respond to a therapeutic vaccine peptide or nucleic acid in accordance with the invention. *Ex vivo* CTL or HTL responses to a particular antigen (infectious or tumor-associated antigen) are induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cell (an infected cell or a tumor cell). Transfected dendritic cells may also be used as antigen presenting cells. Alternatively, dendritic cells are transfected, *e.g.*, with a minigene construct in accordance with the invention, in order to elicit immune responses. Minigenes will be discussed in greater detail in a following section.

Vaccine compositions may also be administered *in vivo* in combination with dendritic cell mobilization whereby loading of dendritic cells occurs *in vivo*.

DNA or RNA encoding one or more of the peptides of the invention can also be administered to a patient. This approach is described, for instance, in Wolff *et. al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; WO 98/04720; and in more detail below. Examples of DNA-based delivery technologies include “naked DNA”, facilitated (bupivicaine, polymers, peptide-mediated) delivery, cationic lipid complexes, and particle-mediated (“gene gun”) or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687).

Preferably, the following principles are utilized when selecting an array of epitopes for inclusion in a polyepitopic composition for use in a vaccine, or for selecting

discrete epitopes to be included in a vaccine and/or to be encoded by nucleic acids such as a minigene. Exemplary epitopes that may be utilized in a vaccine to treat or prevent cancer are set out in Tables XXXVII and XXXVIII. It is preferred that each of the following principles are balanced in order to make the selection. The multiple epitopes to be incorporated in a given vaccine composition may be, but need not be, contiguous in sequence in the native antigen from which the epitopes are derived.

1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with tumor clearance. For HLA Class I this includes 3-4 epitopes that come from at least one TAA. For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one TAA (*see e.g.*, Rosenberg *et al.*, *Science* 278:1447-1450). Epitopes from one TAA may be used in combination with epitopes from one or more additional TAAs to produce a vaccine that targets tumors with varying expression patterns of frequently-expressed TAAs as described, *e.g.*, in Example 15.

2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an IC_{50} of 500 nM or less, or for Class II an IC_{50} of 1000 nM or less.

3.) Sufficient supermotif bearing-peptides, or a sufficient array of allele-specific motif-bearing peptides, are selected to give broad population coverage. For example, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess the breadth, or redundancy of, population coverage.

4.) When selecting epitopes from cancer-related antigens it is often preferred to select analogs because the patient may have developed tolerance to the native epitope. When selecting epitopes for infectious disease-related antigens it is preferable to select either native or analoged epitopes. Of particular relevance for infectious disease vaccines (but for cancer-related vaccines as well), are epitopes referred to as "nested epitopes." Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A peptide comprising "transcendent nested epitopes" is a peptide that has both HLA class I and HLA class II epitopes in it.

When providing nested epitopes, it is preferable to provide a sequence that has the greatest number of epitopes per provided sequence. Preferably, one avoids providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a

longer peptide sequence, such as a sequence comprising nested epitopes, it is important to screen the sequence in order to insure that it does not have pathological or other deleterious biological properties.

5.) When creating a minigene, as disclosed in greater detail in the following section, an objective is to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same as those employed when selecting a peptide comprising nested epitopes. Furthermore, upon determination of the nucleic acid sequence to be provided as a minigene, the peptide encoded thereby is analyzed to determine whether any "junctional epitopes" have been created. A junctional epitope is a potential HLA binding epitope, as predicted, *e.g.*, by motif analysis, that only exists because two discrete peptide sequences are encoded directly next to each other. Junctional epitopes are generally to be avoided because the recipient may bind to an HLA molecule and generate an immune response to that non-native epitope. Of particular concern is a junctional epitope that is a "dominant epitope." A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

IV.K.1. Minigene Vaccines

A growing body of experimental evidence demonstrates that a number of different approaches are available which allow simultaneous delivery of multiple epitopes. Nucleic acids encoding the peptides of the invention are a particularly useful embodiment of the invention. Epitopes for inclusion in a minigene are preferably selected according to the guidelines set forth in the previous section. A preferred means of administering nucleic acids encoding the peptides of the invention uses minigene constructs encoding a peptide comprising one or multiple epitopes of the invention. The use of multi-epitope minigenes is described below and in, *e.g.*, co-pending application U.S.S.N. 09/311,784; Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a multi-epitope DNA plasmid encoding supermotif- and/or motif-bearing CEA epitopes derived from multiple regions of CEA, the PADRE™ universal helper T cell epitope (or multiple HTL epitopes from CEA), and an endoplasmic reticulum-translocating signal sequence

can be engineered. A vaccine may also comprise epitopes, in addition to CEA epitopes, that are derived from other TAAs.

The immunogenicity of a multi-epitopic minigene can be tested in transgenic mice to evaluate the magnitude of CTL induction responses against the epitopes tested. Further, the immunogenicity of DNA-encoded epitopes *in vivo* can be correlated with the *in vitro* responses of specific CTL lines against target cells transfected with the DNA plasmid. Thus, these experiments can show that the minigene serves to both: 1.) generate a CTL response and 2.) that the induced CTLs recognized cells expressing the encoded epitopes.

For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that when translated, a continuous polypeptide sequence is created. To optimize expression and/or immunogenicity, additional elements can be incorporated into the minigene design. Examples of amino acid sequences that can be reverse translated and included in the minigene sequence include: HLA class I epitopes, HLA class II epitopes, a ubiquitination signal sequence, and/or an endoplasmic reticulum targeting signal. In addition, HLA presentation of CTL and HTL epitopes may be improved by including synthetic (*e.g.* poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention.

The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope polypeptide, can then be cloned into a desired expression vector.

Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a down-stream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus

(hCMV) promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

Additional vector modifications may be desired to optimize minigene expression and immunogenicity. In some cases, introns are required for efficient gene expression, and one or more synthetic or naturally-occurring introns could be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.

Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate *E. coli* strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping and DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as a master cell bank and a working cell bank.

In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector, outside the minigene coding sequence, if desired to enhance immunogenicity.

In some embodiments, a bi-cistronic expression vector which allows production of both the minigene-encoded epitopes and a second protein (included to enhance or decrease immunogenicity) can be used. Examples of proteins or polypeptides that could beneficially enhance the immune response if co-expressed include cytokines (*e.g.*, IL-2, IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF), costimulatory molecules, or for HTL responses, pan-DR binding proteins (PADRE™, Epimmune, San Diego, CA). Helper (HTL) epitopes can be joined to intracellular targeting signals and expressed separately from expressed CTL epitopes; this allows direction of the HTL epitopes to a cell compartment different than that of the CTL epitopes. If required, this could facilitate more efficient entry of HTL epitopes into the HLA class II pathway, thereby improving HTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF-β) may be beneficial in certain diseases.

Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and grown to saturation in shaker flasks or a bioreactor

according to well known techniques. Plasmid DNA can be purified using standard bioseparation technologies such as solid phase anion-exchange resins supplied by QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffered saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene DNA vaccines, an alternative method for formulating purified plasmid DNA may be desirable. A variety of methods have been described, and new techniques may become available. Cationic lipids, glycolipids, and fusogenic liposomes can also be used in the formulation (see, *e.g.*, as described by WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Pat No. 5,279,833; WO 91/06309; and Felgner, *et al.*, *Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) could also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

Target cell sensitization can be used as a functional assay for expression and HLA class I presentation of minigene-encoded CTL epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is suitable as a target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. Electroporation can be used for "naked" DNA, whereas cationic lipids allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). These cells are then chromium-51 (^{51}Cr) labeled and used as target cells for epitope-specific CTL lines; cytolysis, detected by ^{51}Cr release, indicates both production of, and HLA presentation of, minigene-encoded CTL epitopes. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

In vivo immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (*e.g.*, IM for DNA in PBS, intraperitoneal (IP) for lipid-complexed DNA).

Twenty-one days after immunization, splenocytes are harvested and restimulated for one week in the presence of peptides encoding each epitope being tested. Thereafter, for CTL effector cells, assays are conducted for cytolysis of peptide-loaded, ^{51}Cr-labeled target cells using standard techniques. Lysis of target cells that were sensitized by HLA loaded with peptide epitopes, corresponding to minigene-encoded epitopes, demonstrates DNA vaccine function for *in vivo* induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of DNA are administered. In a further alternative embodiment, DNA can be adhered to particles, such as gold particles.

IV.K.2. Combinations of CTL Peptides with Helper Peptides

Vaccine compositions comprising the peptides of the present invention, or analogs thereof, which have immunostimulatory activity may be modified to provide desired attributes, such as improved serum half-life, or to enhance immunogenicity.

For instance, the ability of a peptide to induce CTL activity can be enhanced by linking the peptide to a sequence which contains at least one epitope that is capable of inducing a T helper cell response. The use of T helper epitopes in conjunction with CTL epitopes to enhance immunogenicity is illustrated, for example, in the co-pending applications U.S.S.N. 08/820,360, U.S.S.N. 08/197,484, and U.S.S.N. 08/464,234.

Particularly preferred CTL epitope/HTL epitope conjugates are linked by a spacer molecule. The spacer is typically comprised of relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, *e.g.*, Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer need not be comprised of the same residues and thus may be a hetero- or homo-oligomer. When present, the spacer will usually be at least one or two residues, more usually three to six residues. Alternatively, the CTL peptide may be linked to the T helper peptide without a spacer.

The CTL peptide epitope may be linked to the T helper peptide epitope either directly or via a spacer either at the amino or carboxy terminus of the CTL peptide. The amino terminus of either the immunogenic peptide or the T helper peptide may be acylated. The HTL peptide epitopes used in the invention can be modified in the same

manner as CTL peptides. For instance, they may be modified to include D-amino acids or be conjugated to other molecules such as lipids, proteins, sugars and the like.

In certain embodiments, the T helper peptide is one that is recognized by T helper cells present in the majority of the population. This can be accomplished by selecting amino acid sequences that bind to many, most, or all of the HLA class II molecules. These are known as "loosely HLA-restricted" or "promiscuous" T helper sequences. Examples of amino acid sequences that are promiscuous include sequences from antigens such as tetanus toxoid at positions 830-843 (QYIKANSKFIGITE), *Plasmodium falciparum* CS protein at positions 378-398 (DIEKKIAKMEKASSVFNVVNS), and Streptococcus 18kD protein at positions 116 (GAVDSILGGVATYGAA). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

Alternatively, it is possible to prepare synthetic peptides capable of stimulating T helper lymphocytes, in a loosely HLA-restricted fashion, using amino acid sequences not found in nature (*see*, *e.g.*, PCT publication WO 95/07707). These synthetic compounds called Pan-DR-binding epitopes (*e.g.*, PADRE™, Epimmune, Inc., San Diego, CA) are designed to most preferrably bind most HLA-DR (human HLA class II) molecules. For instance, a pan-DR-binding epitope peptide having the formula: aKXVWANTLKAAa, where "X" is either cyclohexylalanine, phenylalanine, or tyrosine, and "a" is either D-alanine or L-alanine, has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type. An alternative of a pan-DR binding epitope comprises all "L" natural amino acids and can be provided in the form of nucleic acids that encode the epitope.

HTL peptide epitopes can also be modified to alter their biological properties. For example, peptides comprising HTL epitopes can contain D-amino acids to increase their resistance to proteases and thus extend their serum half-life. Also, the epitope peptides of the invention can be conjugated to other molecules such as lipids, proteins or sugars, or any other synthetic compounds, to increase their biological activity. Specifically, the T helper peptide can be conjugated to one or more palmitic acid chains at either the amino or carboxyl termini.

In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes cytotoxic T lymphocytes. Lipids have been identified as agents capable of priming CTL *in vivo* against viral antigens. For example, palmitic acid residues can be attached to the ε-and α-

amino groups of a lysine residue and then linked, *e.g.*, via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle, incorporated into a liposome, or emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. A particularly effective immunogen comprises palmitic acid attached to ε- and α- amino groups of Lys, which is attached via linkage, *e.g.*, Ser-Ser, to the amino terminus of the immunogenic peptide.

As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glycerylcysteinlyseryl- serine (P_3CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide (*see, e.g.*, Deres, *et al.*, *Nature* 342:561, 1989). Peptides of the invention can be coupled to P_3CSS, for example, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P_3CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses to infection.

As noted herein, additional amino acids can be added to the termini of a peptide to provide for ease of linking peptides one to another, for coupling to a carrier support or larger peptide, for modifying the physical or chemical properties of the peptide or oligopeptide, or the like. Amino acids such as tyrosine, cysteine, lysine, glutamic or aspartic acid, or the like, can be introduced at the C- or N-terminus of the peptide or oligopeptide, particularly class I peptides. However, it is to be noted that modification at the carboxyl terminus of a CTL epitope may, in some cases, alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the natural sequence by being modified by terminal-NH_2 acylation, *e.g.*, by alkanoyl (C_1-C_{20}) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule.

IV.L. Administration of Vaccines for Therapeutic or Prophylactic Purposes

The peptides of the present invention and pharmaceutical and vaccine compositions of the invention are useful for administration to mammals, particularly humans, to treat and/or prevent cancer. Vaccine compositions containing the peptides of the invention are administered to a cancer patient or to an individual susceptible to, or

otherwise at risk for, cancer to elicit an immune response against TAAs and thus enhance the patient's own immune response capabilities. In therapeutic applications, peptide and/or nucleic acid compositions are administered to a patient in an amount sufficient to elicit an effective CTL and/or HTL response to the tumor antigen and to cure or at least partially arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, *e.g.*, the particular composition administered, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

The vaccine compositions of the invention may also be used purely as prophylactic agents. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 μg and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg. Dosage values for a human typically range from about 500 μg to about 50,000 μg per 70 kilogram patient. This is followed by boosting dosages of between about 1.0 μg to about 50,000 μg of peptide administered at defined intervals from about four weeks to six months after the initial administration of vaccine. The immunogenicity of the vaccine may be assessed by measuring the specific activity of CTL and HTL obtained from a sample of the patient's blood.

As noted above, peptides comprising CTL and/or HTL epitopes of the invention induce immune responses when presented by HLA molecules and contacted with a CTL or HTL specific for an epitope comprised by the peptide. The manner in which the peptide is contacted with the CTL or HTL is not critical to the invention. For instance, the peptide can be contacted with the CTL or HTL either *in vivo* or *in vitro*. If the contacting occurs *in vivo*, the peptide itself can be administered to the patient, or other vehicles, *e.g.*, DNA vectors encoding one or more peptides, viral vectors encoding the peptide(s), liposomes and the like, can be used, as described herein.

When the peptide is contacted *in vitro*, the vaccinating agent can comprise a population of cells, *e.g.*, peptide-pulsed dendritic cells, or TAA-specific CTLs, which have been induced by pulsing antigen-presenting cells *in vitro* with the peptide. Such a cell population is subsequently administered to a patient in a therapeutically effective dose.

For pharmaceutical compositions, the immunogenic peptides of the invention, or DNA encoding them, are generally administered to an individual already diagnosed with

cancer. The peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences.

For therapeutic use, administration should generally begin at the first diagnosis of cancer. This is followed by boosting doses until at least symptoms are substantially abated and for a period thereafter. The embodiment of the vaccine composition (i.e., including, but not limited to embodiments such as peptide cocktails, polyepitopic polypeptides, minigenes, or TAA-specific CTLs) delivered to the patient may vary according to the stage of the disease. For example, a vaccine comprising TAA-specific CTLs may be more efficacious in killing tumor cells in patients with advanced disease than alternative embodiments.

The vaccine compositions of the invention may also be used therapeutically in combination with treatments such as surgery. An example is a situation in which a patient has undergone surgery to remove a primary tumor and the vaccine is then used to slow or prevent recurrence and/or metastasis.

Where susceptible individuals, *e.g.*, individuals who may be diagnosed as being genetically pre-disposed to developing a particular type of tumor, are identified prior to diagnosis of cancer, the composition can be targeted to them, thus minimizing the need for administration to a larger population.

The dosage for an initial therapeutic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1,000 μg and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg. Dosage values for a human typically range from about 500 μg to about 50,000 μg per 70 kilogram patient. Boosting dosages of between about 1.0 μg to about 50,000 μg of peptide pursuant to a boosting regimen over weeks to months may be administered depending upon the patient's response and condition as determined by measuring the specific activity of CTL and HTL obtained from the patient's blood. The peptides and compositions of the present invention may be employed in serious disease states, that is, life-threatening or potentially life threatening situations. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides in preferred compositions of the invention, it is possible and may be felt desirable by the treating physician to administer substantial excesses of these peptide compositions relative to these stated dosage amounts.

Thus, for treatment of cancer, a representative dose is in the range disclosed above, namely where the lower value is about 1, 5, 50, 500, or 1,000 μg and the higher

value is about 10,000; 20,000; 30,000; or 50,000 μg, preferably from about 500 μg to about 50,000 μg per 70 kilogram patient. Initial doses followed by boosting doses at established intervals, *e.g.*, from four weeks to six months, may be required, possibly for a prolonged period of time to effectively immunize an individual. Administration should continue until at least clinical symptoms or laboratory tests indicate that the tumor has been eliminated or that the tumor cell burden has been substantially reduced and for a period thereafter. The dosages, routes of administration, and dose schedules are adjusted in accordance with methodologies known in the art.

The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, intrathecal, or local administration. Preferably, the pharmaceutical compositions are administered parentally, *e.g.*, intravenously, subcutaneously, intradermally, or intramuscularly. Thus, the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, *e.g.*, water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions, such as pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, *etc.*

The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, *i.e.*, from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, *etc.*, in accordance with the particular mode of administration selected.

A human unit dose form of the peptide composition is typically included in a pharmaceutical composition that comprises a human unit dose of an acceptable carrier, preferably an aqueous carrier, and is administered in a volume of fluid that is known by those of skill in the art to be used for administration of such compositions to humans (*see, e.g.*, Remington's Pharmaceutical Sciences, 17^{th} Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pennsylvania, 1985).

The peptides of the invention may also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or to target selectively to infected cells, as well as to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells, such as monoclonal antibodies which bind to the CD45 antigen, or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide compositions. Liposomes for use in accordance with the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, *e.g.*, liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, *e.g.*, Szoka, *et al.*, *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

For targeting cells of the immune system, a ligand to be incorporated into the liposome can include, *e.g.*, antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, *etc.* in a dose which varies according to, *inter alia*, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, and more preferably at a concentration of 25%-75%.

For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, preferably 1%-10%. The surfactant must, of course, be nontoxic, and preferably soluble in the propellant. Representative of such agents are

the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant. A carrier can also be included, as desired, as with, *e.g.*, lecithin for intranasal delivery.

IV.M. Kits

The peptide and nucleic acid compositions of this invention can be provided in kit form together with instructions for vaccine administration. Typically the kit would include desired peptide compositions in a container, preferably in unit dosage form and instructions for administration. An alternative kit would include a minigene construct with desired nucleic acids of the invention in a container, preferably in unit dosage form together with instructions for administration. Lymphokines such as IL-2 or IL-12 may also be included in the kit. Other kit components that may also be desirable include, for example, a sterile syringe, booster dosages, and other desired excipients.

The invention will be described in greater detail by way of specific examples. The following examples are offered for illustrative purposes, and are not intended to limit the invention in any manner. Those of skill in the art will readily recognize a variety of non-critical parameters that can be changed or modified to yield alternative embodiments in accordance with the invention.

V. EXAMPLES

The following examples illustrate identification, selection, and use of immunogenic Class I and Class II peptide epitopes for inclusion in vaccine compositions.

Example 1. HLA Class I and Class II Binding Assays

The following example of peptide binding to HLA molecules demonstrates quantification of binding affinities of HLA class I and class II peptides. Binding assays can be performed with peptides that are either motif-bearing or not motif-bearing.

Epstein-Barr virus (EBV)-transformed homozygous cell lines, fibroblasts, CIR, or 721.221-transfectants were used as sources of HLA class I molecules. These cells were maintained *in vitro* by culture in RPMI 1640 medium supplemented with 2mM L-

glutamine (GIBCO, Grand Island, NY), 50μM 2-ME, 100μg/ml of streptomycin, 100U/ml of penicillin (Irvine Scientific) and 10% heat-inactivated FCS (Irvine Scientific, Santa Ana, CA). Cells were grown in 225-cm^2 tissue culture flasks or, for large-scale cultures, in roller bottle apparatuses. The specific cell lines routinely used for purification of MHC class I and class II molecules are listed in Table XXIV.

Cell lysates were prepared and HLA molecules purified in accordance with disclosed protocols (Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al.*, *J. Immunol.* 154:247 (1995); Sette, *et al.*, *Mol. Immunol.* 31:813 (1994)). Briefly, cells were lysed at a concentration of 10^8 cells/ml in 50 mM Tris-HCl, pH 8.5, containing 1% Nonidet P-40 (Fluka Biochemika, Buchs, Switzerland), 150 mM NaCl, 5 mM EDTA, and 2 mM PMSF. Lysates were cleared of debris and nuclei by centrifugation at 15,000 x g for 30min.

HLA molecules were purified from lysates by affinity chromatography. Lysates prepared as above were passed twice through two pre-columns of inactivated Sepharose CL4-B and protein A-Sepharose. Next, the lysate was passed over a column of Sepharose CL-4B beads coupled to an appropriate antibody. The antibodies used for the extraction of HLA from cell lysates are listed in Table XXV. The anti-HLA column was then washed with 10-column volumes of 10mM Tris-HCL, pH 8.0, in 1% NP-40, PBS, 2-column volumes of PBS, and 2-column volumes of PBS containing 0.4% n-octylglucoside. Finally, MHC molecules were eluted with 50mM diethylamine in 0.15M NaCl containing 0.4% n-octylglucoside, pH 11.5. A 1/25 volume of 2.0M Tris, pH 6.8, was added to the eluate to reduce the pH to ~8.0. Eluates were then concentrated by centrifugation in Centriprep 30 concentrators at 2000 rpm (Amicon, Beverly, MA). Protein content was evaluated by a BCA protein assay (Pierce Chemical Co., Rockford, IL) and confirmed by SDS-PAGE.

A detailed description of the protocol utilized to measure the binding of peptides to Class I and Class II MHC has been published (Sette *et al.*, *Mol. Immunol.* 31:813, 1994; Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998). Briefly, purified MHC molecules (5 to 500nM) were incubated with various unlabeled peptide inhibitors and 1-10nM ^{125}I-radiolabeled probe peptides for 48h in PBS containing 0.05% Nonidet P-40 (NP40) (or 20% w/v digitonin for H-2 IA assays) in the presence of a protease inhibitor cocktail. The final concentrations of protease inhibitors (each from CalBioChem, La Jolla, CA) were 1 mM

PMSF, 1.3 nM 1.10 phenanthroline, 73 μM pepstatin A, 8mM EDTA, 6mM N-ethylmaleimide (for Class II assays), and 200 μM N alpha-p-tosyl-L-lysine chloromethyl ketone (TLCK). All assays were performed at pH 7.0 with the exception of DRB1*0301, which was performed at pH 4.5, and DRB1*1601 (DR2w21β_1) and DRB4*0101 (DRw53), which were performed at pH 5.0. pH was adjusted as described elsewhere (see Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998).

Following incubation, MHC-peptide complexes were separated from free peptide by gel filtration on 7.8 mm x 15 cm TSK200 columns (TosoHaas 16215, Montgomeryville, PA), eluted at 1.2 mls/min with PBS pH 6.5 containing 0.5% NP40 and 0.1% NaN_3. Because the large size of the radiolabeled peptide used for the DRB1*1501 (DR2w2β_1) assay makes separation of bound from unbound peaks more difficult under these conditions, all DRB1*1501 (DR2w2β_1) assays were performed using a 7.8mm x 30cm TSK2000 column eluted at 0.6 mls/min. The eluate from the TSK columns was passed through a Beckman 170 radioisotope detector, and radioactivity was plotted and integrated using a Hewlett-Packard 3396A integrator, and the fraction of peptide bound was determined.

Radiolabeled peptides were iodinated using the chloramine-T method. Representative radiolabeled probe peptides utilized in each assay, and its assay specific IC_{50} nM, are summarized in Tables IV and V. Typically, in preliminary experiments, each MHC preparation was titered in the presence of fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of the total radioactivity. All subsequent inhibition and direct binding assays were performed using these HLA concentrations.

Since under these conditions [label]<[HLA] and $IC_{50} \geq$[HLA], the measured IC_{50} values are reasonable approximations of the true K_D values. Peptide inhibitors are typically tested at concentrations ranging from 120 μg/ml to 1.2 ng/ml, and are tested in two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the IC_{50} of a positive control for inhibition by the IC_{50} for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For database purposes, and inter-experiment comparisons, relative binding values are compiled. These values can subsequently be converted back into IC_{50} nM values by dividing the IC_{50} nM of the

positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation has proven to be the most accurate and consistent for comparing peptides that have been tested on different days, or with different lots of purified MHC.

Because the antibody used for HLA-DR purification (LB3.1) is α-chain specific, β_1 molecules are not separated from β_3 (and/or β_4 and β_5) molecules. The β_1 specificity of the binding assay is obvious in the cases of DRB1*0101 (DR1), DRB1*0802 (DR8w2), and DRB1*0803 (DR8w3), where no β_3 is expressed. It has also been demonstrated for DRB1*0301 (DR3) and DRB3*0101 (DR52a), DRB1*0401 (DR4w4), DRB1*0404 (DR4w14), DRB1*0405 (DR4w15), DRB1*1101 (DR5), DRB1*1201 (DR5w12), DRB1*1302 (DR6w19) and DRB1*0701 (DR7). The problem of β chain specificity for DRB1*1501 (DR2w2β_1), DRB5*0101 (DR2w2β_2), DRB1*1601 (DR2w21β_1), DRB5*0201 (DR51Dw21), and DRB4*0101 (DRw53) assays is circumvented by the use of fibroblasts. Development and validation of assays with regard to DRβ molecule specificity have been described previously (*see, e.g.*, Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998).

Binding assays as outlined above may be used to analyze supermotif and/or motif-bearing epitopes as, for example, described in Example 2.

Example 2. Identification of HLA Supermotif- and Motif-Bearing CTL Candidate Epitopes

Vaccine compositions of the invention may include multiple epitopes that comprise multiple HLA supermotifs or motifs to achieve broad population coverage. This example illustrates the identification of supermotif- and motif-bearing epitopes for the inclusion in such a vaccine composition. Calculation of population coverage is performed using the strategy described below.

Computer searches and algorthims for identification of supermotif and/or motif-bearing epitopes

The searches performed to identify the motif-bearing peptide sequences in Examples 2 and 5 employed protein sequence data for the tumor-associated antigen CEA (GenBank access number M59255).

Computer searches for epitopes bearing HLA Class I or Class II supermotifs or motifs were performed as follows. All translated protein sequences were analyzed using a text string search software program, *e.g.*, MotifSearch 1.4 (D. Brown, San Diego) to identify potential peptide sequences containing appropriate HLA binding motifs; alternative programs are readily produced in accordance with information in the art in view of the motif/supermotif disclosure herein. Furthermore, such calculations can be made mentally. Identified A2-, A3-, and DR-supermotif sequences were scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class I or Class II molecules. These polynomial algorithms take into account both extended and refined motifs (that is, to account for the impact of different amino acids at different positions), and are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$\text{“}\Delta G\text{"} = a_{1i} \times a_{2i} \times a_{3i} \ldots\ldots \times a_{ni}$$

where a_{ji} is a coefficient which represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. The crucial assumption of this method is that the effects at each position are essentially independent of each other (i.e., independent binding of individual side-chains). When residue j occurs at position i in the peptide, it is assumed to contribute a constant amount j_i to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide. This assumption is justified by studies from our laboratories that demonstrated that peptides are bound to MHC and recognized by T cells in essentially an extended conformation (data omitted herein).

The method of derivation of specific algorithm coefficients has been described in Gulukota *et al.*, *J. Mol. Biol.* 267:1258-126, 1997; (*see* also Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; and Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). Briefly, for all i positions, anchor and non-anchor alike, the geometric mean of the average relative binding (ARB) of all peptides carrying j is calculated relative to the remainder of the group, and used as the estimate of j_i. For Class II peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure. To calculate an algorithm score of a given peptide in a test set, the ARB values corresponding to the sequence of the peptide are multiplied. If this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

Selection of HLA-A2 supertype cross-reactive peptides

The complete protein sequence from CEA was scanned, utilizing motif identification software, to identify 8-, 9-, 10-, and 11-mer sequences containing the HLA-A2-supermotif main anchor specificity.

A total of 336 HLA-A2 supermotif-positive sequences were identified. Of these, 266 peptides corresponding to the sequences were then synthesized and tested for their capacity to bind purified HLA-A*0201 molecules *in vitro* (HLA-A*0201 is considered a prototype A2 supertype molecule). Fourteen of the 266 peptides bound A*0201 with IC_{50} values ≤500 nM.

The fourteen A*0201-binding peptides were subsequently tested for the capacity to bind to additional A2-supertype molecules (A*0202, A*0203, A*0206, and A*6802). As shown in Table XXVI, 10 of the 14 peptides were found to be A2-supertype cross-reactive binders, binding at least three of the five A2-supertype alleles tested.

Selection of HLA-A3 supermotif-bearing epitopes

The protein sequences scanned above are also examined for the presence of peptides with the HLA-A3-supermotif primary anchors using methodology similar to that performed to identify HLA-A2 supermotif-bearing epitopes.

Peptides corresponding to the supermotif-bearing sequences are then synthesized and tested for binding to HLA-A*0301 and HLA-A*1101 molecules, the two most prevalent A3-supertype alleles. The peptides that are found to bind one of the two alleles with binding affinities of ≤500 nM are then tested for binding cross-reactivity to the other common A3-supertype alleles (A*3101, A*3301, and A*6801) to identify those that can bind at least three of the five HLA-A3-supertype molecules tested.

Selection of HLA-B7 supermotif bearing epitopes

The same target antigen protein sequences are also analyzed to identify HLA-B7-supermotif-bearing sequences. The corresponding peptides are then synthesized and tested for binding to HLA-B*0702, the most common B7-supertype allele (*i.e.*, the prototype B7 supertype allele). Those peptides that bind B*0702 with IC_{50} of ≤500 nM are then tested for binding to other common B7-supertype molecules (B*3501, B*5101,

B*5301, and B*5401) to identify those peptides that are capable of binding to three or more of the five B7-supertype alleles tested.

Selection of A1 and A24 motif-bearing epitopes

To further increase population coverage, HLA-A1 and -A24 epitopes can also be incorporated into potential vaccine constructs. An analysis of the protein sequence data from the target antigens utilized above can also be performed to identify HLA-A1- and A24-motif-containing conserved sequences.

Example 3. Confirmation of Immunogenicity

Nine of the ten cross-reactive candidate CTL A2-supermotif-bearing peptides were selected for *in vitro* immunogenicity testing. Testing was performed using the following methodology:

Target Cell Lines for Cellular Screening:

The .221A2.1 cell line, produced by transferring the HLA-A2.1 gene into the HLA-A, -B, -C null mutant human B-lymphoblastoid cell line 721.221, was used as the peptide-loaded target to measure activity of HLA-A2.1-restricted CTL. The HLA-typed melanoma cell lines (624mel and 888mel) were obtained from Y. Kawakami and S. Rosenberg, National Cancer Institute, Bethesda, MD. The colon adenocarcinoma cell lines SW403 and HT-20, the osteosarcoma line Saos-2 and the breast tumor line BT540 were obtained from the American Type Culture Collection (ATCC) (Rockville, MD). The gastric cancer line, KATO III was obtained from the Japanese Cancer Research Resources Bank. The Saos-2/175 (Saos-2 transfected with the p53 gene containing a mutation at position 175) was obtained from Dr. Levine, Princeton University, Princeton, NJ. The cell lines that were obtained from ATCC were maintained under the culture conditions recommended by the supplier. All other cell lines were grown in RPMI-1640 medium supplemented with antibiotics, sodium pyruvate, nonessential amino acids and 10% (v/v) heat inactivated FCS. The melanoma, colon and gastric cancer cells were treated with 100U/ml IFNγ (Genzyme) for 48 hours at 37°C before use as targets in the ^{51}Cr release and *in situ* IFNγ assays. The p53 tumor targets were treated with 20 ng/ml IFNγ and 3 ng/ml TNFα for 24 hours prior to assay (*see, e.g.*, Theobald *et al.*, *Proc. Natl. Acad. Sci. USA* 92:11993, 1995).

Primary CTL Induction Cultures:

Generation of Dendritic Cells (DC): PBMCs were thawed in RPMI with 30 μg/ml DNAse, washed twice and resuspended in complete medium (RPMI-1640 plus 5% AB human serum, non-essential amino acids, sodium pyruvate, L-glutamine and penicillin/strpetomycin). The monocytes were purified by plating 10 x 10^6 PBMC/well in a 6-well plate. After 2 hours at 37°C, the non-adherent cells were removed by gently shaking the plates and aspirating the supernatants. The wells were washed a total of three times with 3 ml RPMI to remove most of the non-adherent and loosely adherent cells. Three ml of complete medium containing 50 ng/ml of GM-CSF and 1,000 U/ml of IL-4 were then added to each well. DC were used for CTL induction cultures following 7 days of culture.

Induction of CTL with DC and Peptide: CD8+ T-cells were isolated by positive selection with Dynal immunomagnetic beads (Dynabeads® M-450) and the detacha-bead® reagent. Typically about 200-250x10^6 PBMC were processed to obtain 24x10^6 $CD8^+$ T-cells (enough for a 48-well plate culture). Briefly, the PBMCs were thawed in RPMI with 30μg/ml DNAse, washed once with PBS containing 1% human AB serum and resuspended in PBS/1% AB serum at a concentration of 20x10^6cells/ml. The magnetic beads were washed 3 times with PBS/AB serum, added to the cells (140μl beads/20x10^6 cells) and incubated for 1 hour at 4°C with continuous mixing. The beads and cells were washed 4x with PBS/AB serum to remove the nonadherent cells and resuspended at 100x10^6 cells/ml (based on the original cell number) in PBS/AB serum containing 100μl/ml detacha-bead® reagent and 30μg/ml DNAse. The mixture is incubated for 1 hour at room temperature with continuous mixing. The beads were washed again with PBS/AB/DNAse to collect the CD8+ T-cells. The DC were collected and centrifuged at 1300 rpm for 5-7 minutes, washed once with PBS with 1% BSA, counted and pulsed with 40μg/ml of peptide at a cell concentration of 1-2x10^6/ml in the presence of 3μg/ml $ß_2$- microglobulin for 4 hours at 20°C. The DC were then irradiated (4,200 rads), washed 1 time with medium and counted again.

Setting up induction cultures: 0.25 ml cytokine-generated DC (@1x10^5 cells/ml) were co-cultured with 0.25ml of CD8+ T-cells (@2x10^6 cell/ml) in each well of a 48-well plate in the presence of 10 ng/ml of IL-7. rHuman IL10 was added the next day at a final concentration of 10 ng/ml and rhuman IL2 was added 48 hours later at 10IU/ml.

Restimulation of the induction cultures with peptide-pulsed adherent cells: Seven and fourteen days after the primary induction the cells were restimulated with peptide-pulsed adherent cells. The PBMCS were thawed and washed twice with RPMI and DNAse. The cells were resuspended at $5x10^6$ cells/ml and irradiated at ~4200 rads. The PBMCs were plated at $2x10^6$ in 0.5ml complete medium per well and incubated for 2 hours at 37°C. The plates were washed twice with RPMI by tapping the plate gently to remove the nonadherent cells and the adherent cells pulsed with 10μg/ml of peptide in the presence of 3 μg/ml $ß_2$ microglobulin in 0.25ml RPMI/5%AB per well for 2 hours at 37°C. Peptide solution from each well was aspirated and the wells were washed once with RPMI. Most of the media was aspirated from the induction cultures (CD8+ cells) and brought to 0.5 ml with fresh media. The cells were then transferred to the wells containing the peptide-pulsed adherent cells. Twenty four hours later rhuman IL10 was added at a final concentration of 10ng/ml and rhuman IL2 was added the next day and again 2-3 days later at 50IU/ml (Tsai *et al.*, *Critical Reviews in Immunology* 18(1-2):65-75, 1998). Seven days later the cultures were assayed for CTL activity in a ^{51}Cr release assay. In some experiments the cultures were assayed for peptide-specific recognition in the in situ IFNγ ELISA at the time of the second restimulation followed by assay of endogenous recognition 7 days later. After expansion, activity was measured in both assays for a side by side comparison.

Measurement of CTL lytic activity by ^{51}Cr release.

Seven days after the second restimulation, cytotoxicity was determined in a standard (5hr) ^{51}Cr release assay by assaying individual wells at a single E:T. Peptide-pulsed targets were prepared by incubating the cells with 10μg/ml peptide overnight at 37°C.

Adherent target cells were removed from culture flasks with trypsin-EDTA. Target cells were labelled with 200μCi of ^{51}Cr sodium chromate (Dupont, Wilmington, DE) for 1 hour at 37°C. Labelled target cells are resuspended at 10^6 per ml and diluted 1:10 with K562 cells at a concentration of $3.3x10^6$/ml (an NK-sensitive erythroblastoma cell line used to reduce non-specific lysis). Target cells (100 μl) and 100μl of effectors were plated in 96 well round-bottom plates and incubated for 5 hours at 37°C. At that time, 100 μl of supernatant were collected from each well and percent lysis was determined according to the formula: [(cpm of the test sample- cpm of the spontaneous ^{51}Cr release sample)/(cpm of the maximal ^{51}Cr release sample- cpm of the

spontaneous ^{51}Cr release sample)] x 100. Maximum and spontaneous release were determined by incubating the labelled targets with 1% Trition X-100 and media alone, respectively. A positive culture was defined as one in which the specific lysis (sample-background) was 10% or higher in the case of individual wells and was 15% or more at the 2 highest E:T ratios when expanded cultures were assayed.

***In situ* Measurement of Human γIFN Production as an Indicator of Peptide-specific and Endogenous Recognition**

Immulon 2 plates were coated with mouse anti-human IFNγ monoclonal antibody (4 μg/ml 0.1M $NaHCO_3$, pH8.2) overnight at 4°C. The plates were washed with Ca^{2+}, Mg^{2+}-free PBS/0.05% Tween 20 and blocked with PBS/10% FCS for 2 hours, after which the CTLs (100 μl/well) and targets (100 μl/well) were added to each well, leaving empty wells for the standards and blanks (which received media only). The target cells, either peptide-pulsed or endogenous targets, were used at a concentration of $1x10^6$ cells/ml. The plates were incubated for 48 hours at 37°C with 5% CO_2.

Recombinant human IFNγ was added to the standard wells starting at 400 pg or 1200pg/100μl/well and the plate incubated for 2 hours at 37°C. The plates were washed and 100 μl of biotinylated mouse anti-human IFNγ monoclonal antibody (4μg/ml in PBS/3%FCS/0.05% Tween 20) were added and incubated for 2 hours at room temperature. After washing again, 100 μl HRP-streptavidin were added and incubated for 1 hour at room temperature. The plates were then washed 6x with wash buffer, 100μl/well developing solution (TMB 1:1) were added, and the plates allowed to develop for 5-15 minutes. The reaction was stopped with 50 μl/well 1M H_3PO_4 and read at OD450. A culture was considered positive if it measured at least 50 pg of IFNγ/well above background and was twice the background level of expression.

CTL Expansion. Those cultures that demonstrated specific lytic activity against peptide-pulsed targets and/or tumor targets were expanded over a two week period with anti-CD3. Briefly, $5x10^4$ CD8+ cells were added to a T25 flask containing the following: $1x10^6$ irradiated (4,200 rad) PBMC (autologous or allogeneic) per ml, $2x10^5$ irradiated (8,000 rad) EBV- transformed cells per ml, and OKT3 (anti-CD3) at 30ng per ml in RPMI-1640 containing 10% (v/v) human AB serum, non-essential amino acids, sodium pyruvate, 25μM 2-mercaptoethanol, L-glutamine and penicillin/streptomycin. rHuman IL2 was added 24 hours later at a final concentration of

200IU/ml and every 3 days thereafter with fresh media at 50IU/ml. The cells were split if the cell concentration exceeded $1x10^6$/ml and the cultures were assayed between days 13 and 15 at E:T ratios of 30, 10, 3 and 1:1 in the ^{51}Cr release assay or at $1x10^6$/ml in the *in situ* IFNγ assay using the same targets as before the expansion.

Immunogenicity of A2 supermotif-bearing peptides

Nine of the ten A2-supermotif cross-reactive binding peptides were tested in the cellular assay for the ability to induce peptide-specific CTL in normal individuals. In this analysis, a peptide was considered to be an epitope if it induced peptide-specific CTLs in at least 2 donors (unless otherwise noted) and if those CTLs also recognized the endogenously expressed peptide. Of these nine, six were able to induce a peptide-specific CTL response in at least 2 normal donors. Further analysis demonstrated that 5 of these also recognized target cells pulsed with the wild-type peptide and tumor targets that endogenously express CEA (Table XXVII).

The CEA epitopes 691 and 605 were previously identifed (*see* Kawashima *et al.*, *Hum. Immunol.* 59:1-14, 1998). The other four immunogenic epitopes were further evaluated. Peptide specific CTLs to CEA.233, CEA.569, and CEA.687 were observed in one to two donors but endogenous recognition was observed only with CEA.687.

The CTL that demonstrated a positive response to CEA.687 in a ^{51}Cr release assay were expanded and re-assayed against peptide-pulsed and endogenous target. Of the four individual cultures, three also recognized the endogenous target. One culture demonstrated significant lysis of peptide-pulsed target, but not tumor target. Two of the individual positive cultures were also tested against 221A2.1 target cells pulsed with different concentrations of peptide to measure CTL avidity. One line demonstrated high specific lysis at concentrations down to 1 ng/ml while both cultures exhibited a titration of activity further validating CEA.687 as an epitope. In a cold target inhibition assay in which peptide-pulsed targets were incubated with ^{51}Cr-labelled targets to compete for lysis by the CTL, lysis of radiolabelled target cells by two different CTL lines was blocked by increasing the number of target cells pulsed with CEA.687. The non-specific peptide HBVc.18 did not inhibit lysis, thus further demonstrating the epitope specificity of the CTLs.

*Evaluation of A*03/A11 immunogenicity*

HLA-A3 supermotif-bearing cross-reactive binding peptides are also evaluated for immunogenicity using methodology analogous for that used to evaluate the immunogenicity of the HLA-A2 supermotif peptides.

Evaluation of B7 immunogenicity

Immunogenicity screening of the B7-supertype cross-reactive binding peptides identified in Example 2 are evaluated in a manner analogous to the evaluation of A2-and A3-supermotif-bearing peptides.

Example 4. Implementation of the Extended Supermotif to Improve the Binding Capacity of Native Epitopes by Creating Analogs

HLA motifs and supermotifs (comprising primary and/or secondary residues) are useful in the identification and preparation of highly cross-reactive native peptides, as demonstrated herein. Moreover, the definition of HLA motifs and supermotifs also allows one to engineer highly cross-reactive epitopes by identifying residues within a native peptide sequence which can be analogued, or "fixed" to confer upon the peptide certain characteristics, *e.g.* greater cross-reactivity within the group of HLA molecules that comprise a supertype, and/or greater binding affinity for some or all of those HLA molecules. Examples of analog peptides that exhibit modulated binding affinity are set forth in this example.

Analoguing at Primary Anchor Residues

Peptide engineering strategies were implemented to further increase the cross-reactivity of the epitopes identified above. On the basis of the data disclosed, *e.g.*, in related and co-pending U.S.S.N 09/226,775, the main anchors of A2-supermotif-bearing peptides are altered, for example, to introduce a preferred L, I, V, or M at position 2, and I or V at the C-terminus.

Peptides that exhibit at least weak A*0201 binding (IC_{50} of 5000 nM or less), and carrying suboptimal anchor residues at either position 2, the C-terminal position, or both, can be fixed by introducing canonical substitutions (L at position 2 and V at the C-terminus). Those analogued peptides that show at least a three-fold increase in A*0201 binding and bind with an IC_{50} of 500 nM, or less were then tested for A2 cross-reactive binding along with their wild-type (WT) counterparts. Analogued peptides that bind at

least three of the five A2 supertype alleles were then selected for cellular screening analysis.

Additionally, the selection of analogs for cellular screening analysis was further restricted by the capacity of the WT parent peptide to bind at least weakly, *i.e.*, bind at an IC_{50} of 5000nM or less, to three of more A2 supertype alleles. The rationale for this requirement is that the WT peptides must be present endogenously in sufficient quantity to be biologically relevant. Analogued peptides have been shown to have increased immunogenicity and cross-reactivity by T cells specific for the WT epitope (*see*, *e.g.*, Parkhurst *et al.*, *J. Immunol.* 157:2539, 1996; and Pogue *et al.*, *Proc. Natl. Acad. Sci. USA* 92:8166, 1995).

In the cellular screening of these peptide analogs, it is important to demonstrate that analog-specific CTLs are also able to recognize the wild-type peptide and, when possible, tumor targets that endogenously express the epitope.

Sixty-five CEA peptides met the criteria for analoguing at primary anchor residues by introducing a canonical substitution: these peptides showed at least weak A*0201 binding (IC_{50} of 5000 nM or less) and carried suboptimal anchor residues.

Ten analogs of nine of these peptides were generated and evaluated for cross-reactive binding to other A2 supertype molecules (Table XXII). Eight of these bound minimally to 3 of the 5 A2 supertype alleles, and their WT parents also bound at least weakly to 3 of 5 alleles. In the case of peptide CEA.605, the analog did not exhibit a three-fold increase in A*0201 binding affinity. This peptide did, however, show increased cross-reactivity and therefore was included in the selection of peptides to be analyzed for immunogenicity.

Eight analogs were selected for cellular screening studies. One of these CEA.24V9, was previously identified as an epitope (Kawashima *et al.*, *Hum. Immunol.* 59:1-14, 1998). Three additional peptides were screened and, as shown in Table XXVIII, CEA.233V10, CEA.605V9, and CEA.589V9 all induced CTL that were able to recognize peptide-pulsed and/or tumor targets. After expansion of the positive cultures, the CTLs were again tested against the analog and the parental WT peptide and tumor targets. CTLs to both analogs demonstrated recognition of the WT peptide and the tumor cell line, KATO III. In addition to being immunogenic, CEA.233V10 and CEA.605V9 showed improved overall binding when compared to the corresponding WT peptide as well as cross-reactive binding to 4 alleles. An additional epitope, CEA.589V9, was

immunogenic and CEA.589V9-specific CTLs recognized the wildtype peptide, but endogenous recognition was not observed.

Using methodology similar to that used to develop HLA-A2 analogs, analogs of HLA-A3 and HLA-B7 supermotif-bearing epitopes may also be generated. For example, peptides binding at least weakly to 3/5 of the A3-supertype molecules may be engineered at primary anchor residues to possess a preferred residue (V, S, M, or A) at position 2. The analog peptides are then tested for the ability to bind A*03 and A*11 (prototype A3 supertype alleles). Those peptides that demonstrate ≤ 500 nM binding capacity are then tested for A3-supertype cross-reactivity. B7 supermotif-bearing peptides may, for example, be engineered to possess a preferred residue (V, I, L, or F) at the C-terminal primary anchor position, as demonstrated by Sidney *et al.* (*J. Immunol.* 157:3480-3490, 1996) and tested for binding to B7 supertype alleles.

Analoguing at Secondary Anchor Residues

Moreover, HLA supermotifs are of value in engineering highly cross-reactive peptides and/or peptides that bind HLA molecules with increased affinity by identifying particular residues at secondary anchor positions that are associated with such properties. For example, the binding capacity of a B7 supermotif-bearing peptide representing a discreet single amino acid substitution at position 1 can be analyzed. A peptide can, for example, be analogued to substitute L with F at position 1 and subsequently be evaluated for increased binding affinity/ and or increased cross-reactivity. This procedure will identify analogued peptides with modulated binding affinity.

Engineered analogs with sufficiently improved binding capacity or cross-reactivity are tested for immunogenicity as above.

Other analoguing strategies

Another form of peptide analoguing, unrelated to the anchor positions, involves the substitution of a cysteine with α-amino butyric acid. Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Subtitution of α-amino butyric acid for cysteine not only alleviates this problem, but has been shown to improve binding and crossbinding capabilities in some instances (*see, e.g.*, the review by Sette *et al.*, In:

Persistent Viral Infections, Eds. R. Ahmed and I. Chen, John Wiley & Sons, England, 1999).

In conclusion, these data demonstrate that by the use of even single amino acid substitutions, it is possible to increase the binding affinity and/or cross-reactivity of peptide ligands for HLA supertype molecules.

Example 5. Identification of peptide epitope sequences with HLA-DR binding motifs

Peptide epitopes bearing an HLA class II supermotif or motif may also be identified as outlined below using methodology similar to that described in Examples 1-3.

Selection of HLA-DR-supermotif-bearing epitopes

To identify HLA class II HTL epitopes, the CEA protein sequence was analyzed for the presence of sequences bearing an HLA-DR-motif or supermotif. Specifically, 15-mer sequences were selected comprising a DR-supermotif, further comprising a 9-mer core, and three-residue N- and C-terminal flanking regions (15 amino acids total).

Protocols for predicting peptide binding to DR molecules have been developed (Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). These protocols, specific for individual DR molecules, allow the scoring, and ranking, of 9-mer core regions. Each protocol not only scores peptide sequences for the presence of DR-supermotif primary anchors (i.e., at position 1 and position 6) within a 9-mer core, but additionally evaluates sequences for the presence of secondary anchors. Using allele specific selection tables (see, *e.g.*, Southwood *et al.*, *ibid.*), it has been found that these protocols efficiently select peptide sequences with a high probability of binding a particular DR molecule. Additionally, it has been found that performing these protocols in tandem, specifically those for DR1, DR4w4, and DR7, can efficiently select DR cross-reactive peptides.

The CEA-derived peptides identified above were tested for their binding capacity for various common HLA-DR molecules. All peptides were initially tested for binding to the DR molecules in the primary panel: DR1, DR4w4, and DR7. Peptides binding at least 2 of these 3 DR molecules with an IC_{50} value of 1000 nM or less, were then tested for binding to DR5*0101, DRB1*1501, DRB1*1101, DRB1*0802, and DRB1*1302. Peptides were considered to be cross-reactive DR supertype binders if they bound at an IC_{50} value of 1000 nM or less to at least 5 of the 8 alleles tested.

Following the strategy outlined above, 100 DR supermotif-bearing sequences were identified within the CEA protein sequence. Of those, 24 scored positive in 2 of the

3 combined DR 147 algorithms. These peptides were synthesized and tested for binding to HLA-DRB1*0101, DRB1*0401, DRB1*0701. Of the 24 peptides tested, 10 bound at least 2 of the 3 alleles (Table XXIX).

These 10 peptides were then tested for binding to secondary DR supertype alleles: DRB5*0101, DRB1*1501, DRB1*1101, DRB1*0802, and DRB1*1302. Five peptides were identified that bound at least 5 of the 8 alleles tested and which occurred in distinct, non-overlapping regions (Table XXX).

Selection of DR3 motif peptides

Because HLA-DR3 is an allele that is prevalent in Caucasian, Black, and Hispanic populations, DR3 binding capacity is an important criterion in the selection of HTL epitopes. However, data generated previously indicated that DR3 only rarely cross-reacts with other DR alleles (Sidney *et al., J. Immunol.* 149:2634-2640, 1992; Geluk *et al., J. Immunol.* 152:5742-5748, 1994; Southwood *et al., J. Immunol.* 160:3363-3373, 1998). This is not entirely surprising in that the DR3 peptide-binding motif appears to be distinct from the specificity of most other DR alleles. For maximum efficiency in developing vaccine candidates it would be desirable for DR3 motifs to be clustered in proximity with DR supermotif regions. Thus, peptides shown to be candidates may also be assayed for their DR3 binding capacity. However, in view of the distinct binding specificity of the DR3 motif, peptides binding only to DR3 can also be considered as candidates for inclusion in a vaccine formulation.

To efficiently identify peptides that bind DR3, the CEA protein sequence was analyzed for conserved sequences carrying one of the two DR3 specific binding motifs (Table III) reported by Geluk *et al.* (*J. Immunol.* 152:5742-5748, 1994). Thirty motif-positive peptides were identified. The corresponding peptides were then synthesized and tested for the ability to bind DR3 with an affinity of 1000 nM or better, *i.e.*, less than 1000 nM. Two peptides were found that met this binding criterion (Table XXXI), and thereby qualify as HLA class II high affinity binders. Additionally, the 2 DR3 binders were tested for binding to the DR supertype alleles (Table XXXII). For both peptides, binding to other DR supertype molecules was observed, but neither peptide could be categorized as a DR supertype cross-reactive binding peptide. Conversely, the DR supertype cross-reactive binding peptides were also tested for DR3 binding capacity. One peptide, CEA.50, exhibited DR3 binding (Table XXXII).

DR3 binding epitopes identified in this manner may then be included in vaccine compositions with DR supermotif-bearing peptide epitopes.

In summary, 5 DR supertype cross-reactive binding peptides and 3 DR3 binding peptides were identified from the CEA protein sequence, with one peptide shared between the two motifs.

Example 6. Immunogenicity of HTL epitopes

This example determines immunogenic DR supermotif- and DR3 motif-bearing epitopes among those identified using the methodology in Example 5. Immunogenicity of HTL epitopes are evaluated in a manner analogous to the determination of immunogenicity of CTL epitopes by assessing the ability to stimulate HTL responses and/or by using appropriate transgenic mouse models. Immunogenicity is determined by screening for: 1.) *in vitro* primary induction using normal PBMC or 2.) recall responses from cancer patient PBMCs.

Example 7. Calculation of phenotypic frequencies of HLA-supertypes in various ethnic backgrounds to determine breadth of population coverage

This example illustrates the assessment of the breadth of population coverage of a vaccine composition comprised of multiple epitopes comprising multiple supermotifs and/or motifs.

In order to analyze population coverage, gene frequencies of HLA alleles were determined. Gene frequencies for each HLA allele were calculated from antigen or allele frequencies utilizing the binomial distribution formulae gf=1-(SQRT(1-af)) (see, *e.g.*, Sidney *et al.*, *Human Immunol.* 45:79-93, 1996). To obtain overall phenotypic frequencies, cumulative gene frequencies were calculated, and the cumulative antigen frequencies derived by the use of the inverse formula [af=1-(1-Cgf)^2].

Where frequency data was not available at the level of DNA typing, correspondence to the serologically defined antigen frequencies was assumed. To obtain total potential supertype population coverage no linkage disequilibrium was assumed, and only alleles confirmed to belong to each of the supertypes were included (minimal estimates). Estimates of total potential coverage achieved by inter-loci combinations were made by adding to the A coverage the proportion of the non-A covered population that could be expected to be covered by the B alleles considered (*e.g.*, total=A+B*(1-A)). Confirmed members of the A3-like supertype are A3, A11, A31, A*3301, and A*6801.

Although the A3-like supertype may also include A34, A66, and A*7401, these alleles were not included in overall frequency calculations. Likewise, confirmed members of the A2-like supertype family are A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*6802, and A*6901. Finally, the B7-like supertype-confirmed alleles are: B7, B*3501-03, B51, B*5301, B*5401, B*5501-2, B*5601, B*6701, and B*7801 (potentially also B*1401, B*3504-06, B*4201, and B*5602).

Population coverage achieved by combining the A2-, A3- and B7-supertypes is approximately 86% in five major ethnic groups (see Table XXI). Coverage may be extended by including peptides bearing the A1 and A24 motifs. On average, A1 is present in 12% and A24 in 29% of the population across five different major ethnic groups (Caucasian, North American Black, Chinese, Japanese, and Hispanic). Together, these alleles are represented with an average frequency of 39% in these same ethnic populations. The total coverage across the major ethnicities when A1 and A24 are combined with the coverage of the A2-, A3- and B7-supertype alleles is >95%. An analogous approach can be used to estimate population coverage achieved with combinations of class II motif-bearing epitopes.

Example 8. Recognition Of Endogenous Processed Antigens After Priming

This example determines that CTL induced by native or analogued peptide epitopes identified and selected as described in Examples 1-6 recognize endogenously synthesized, *i.e.*, native antigens, using a transgenic mouse model.

Effector cells isolated from transgenic mice that are immunized with peptide epitopes (as described, e.g., in Wentworth et al., *Mol. Immunol.* 32:603, 1995), for example HLA-A2 supermotif-bearing epitopes, are re-stimulated *in vitro* using peptide-coated stimulator cells. Six days later, effector cells are assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on ^{51}Cr labeled Jurkat-A2.1/K^b target cells in the absence or presence of peptide, and also tested on ^{51}Cr labeled target cells bearing the endogenously synthesized antigen, *i.e.* cells that are stably transfected with TAA expression vectors.

The result will demonstrate that CTL lines obtained from animals primed with peptide epitope recognize endogenously synthesized antigen. The choice of transgenic mouse model to be used for such an analysis depends upon the epitope(s) that is being evaluated. In addition to HLA-A*0201/K^b transgenic mice, several other transgenic

mouse models including mice with human A11, which may also be used to evaluate A3 epitopes, and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed, which may be used to evaluate HTL epitopes.

Example 9. Activity Of CTL-HTL Conjugated Epitopes In Transgenic Mice

This example illustrates the induction of CTLs and HTLs in transgenic mice by use of a tumor associated antigen CTL/HTL peptide conjugate whereby the vaccine composition comprises peptides to be administered to a cancer patient. The peptide composition can comprise multiple CTL and/or HTL epitopes and further, can comprise epitopes selected from multiple-tumor associated antigens. The epitopes are identified using methodology as described in Examples 1-6 This analysis demonstrates the enhanced immunogenicity that can be achieved by inclusion of one or more HTL epitopes in a vaccine composition. Such a peptide composition can comprise an HTL epitope conjugated to a preferred CTL epitope containing, for example, at least one CTL epitope selected from Tables XXII, XXVI, XXVII, or other analogs of that epitope. The HTL epitope is, for example, selected from Table XXXII. The peptides may be lipidated, if desired.

Immunization procedures: Immunization of transgenic mice is performed as described (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). For example, A2/K^b mice, which are transgenic for the human HLA A2.1 allele and are useful for the assessment of the immunogenicity of HLA-A*0201 motif- or HLA-A2 supermotif-bearing epitopes, are primed subcutaneously (base of the tail) with 0.1 ml of peptide conjugate formulated in saline, or DMSO/saline. Seven days after priming, splenocytes obtained from these animals are restimulated with syngenic irradiated LPS-activated lymphoblasts coated with peptide.

The target cells for peptide-specific cytotoxicity assays are Jurkat cells transfected with the HLA-A2.1/K^b chimeric gene (*e.g.*, Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991).

In vitro CTL activation: One week after priming, spleen cells ($30x10^6$ cells/flask) are co-cultured at 37°C with syngeneic, irradiated (3000 rads), peptide coated lymphoblasts ($10x10^6$ cells/flask) in 10 ml of culture medium/T25 flask. After six days, effector cells are harvested and assayed for cytotoxic activity.

Assay for cytotoxic activity: Target cells (1.0 to $1.5x10^6$) are incubated at 37°C in the presence of 200 μl of ^{51}Cr. After 60 minutes, cells are washed three times and

resuspended in medium. Peptide is added where required at a concentration of 1 μg/ml. For the assay, 10^4 ^{51}Cr-labeled target cells are added to different concentrations of effector cells (final volume of 200 μl) in U-bottom 96-well plates. After a 6 hour incubation period at 37°C, a 0.1 ml aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent specific lysis is determined by the formula: percent specific release = 100 x (experimental release - spontaneous release)/(maximum release - spontaneous release). To facilitate comparison between separate CTL assays run under the same conditions, % ^{51}Cr release data is expressed as lytic units/10^6 cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a 6 hour ^{51}Cr release assay. To obtain specific lytic units/10^6, the lytic units/10^6 obtained in the absence of peptide is subtracted from the lytic units/10^6 obtained in the presence of peptide. For example, if 30% ^{51}Cr release is obtained at the effector (E): target (T) ratio of 50:1 (i.e., $5x10^5$ effector cells for 10,000 targets) in the absence of peptide and 5:1 (i.e., $5x10^4$ effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be: $[(1/50,000)-(1/500,000)] \times 10^6 = 18$ LU.

The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation. The frequency and degree of CTL response can also be compared to the CTL response achieved using the CTL epitopes by themselves. Analyses similar to this may be performed to evaluate the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

Example 10. Selection of CTL and HTL epitopes for inclusion in a cancer vaccine.

This example illustrates the procedure for the selection of peptide epitopes for vaccine compositions of the invention. The peptides in the composition may be in the form of a nucleic acid sequence, either single or one or more sequences (*i.e.*, minigene) that encodes peptide(s), or may be single and/or polyepitopic peptides.

The following principles are utilized when selecting an array of epitopes for inclusion in a vaccine composition. Each of the following principles are balanced in order to make the selection.

1.) Epitopes are selected which, upon administration, mimic immune responses that have been observed to be correlated with tumor clearance. For HLA Class I this includes 3-4 epitopes that come from at least one TAA. For HLA Class II a similar rationale is employed; again 3-4 epitopes are selected from at least one TAA (*see e.g.*, Rosenberg *et al.*, *Science* 278:1447-1450). Epitopes from one TAA may be used in combination with epitopes from one or more additional TAAs to produce a vaccine that targets tumors with varying expression patterns of frequently-expressed TAAs as described, *e.g.*, in Example 15.

2.) Epitopes are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an IC_{50} of 500 nM or less, or for Class II an IC_{50} of 1000 nM or less.

3.) Sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. For example, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art and discussed herein, can be employed to assess breadth, or redundancy, of population coverage.

4.) When selecting epitopes from cancer-related antigens it is often preferred to select analogs because the patient may have developed tolerance to the native epitope. When selecting epitopes for infectious disease-related antigens it is preferable to select either native or analoged epitopes. Of relevance for infectious disease vaccines (but for cancer-related vaccines as well), are epitopes referred to as "nested epitopes." Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. A peptide comprising "transcendent nested epitopes" is a peptide that has both HLA class I and HLA class II epitopes in it.

When providing nested epitopes, a sequence that has the greatest number of epitopes per provided sequence is provided. A limitation on this principle is to avoid providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a longer peptide sequence, such as a sequence comprising nested epitopes, the sequence is screened in order to insure that it does not have pathological or other deleterious biological properties.

5.) When creating a minigene, as disclosed in greater detail in Example 11, an objective is to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same as those employed when

selecting a peptide comprising nested epitopes. Additionally, however, upon determination of the nucleic acid sequence to be provided as a minigene, the peptide sequence encoded thereby is analyzed to determine whether any "junctional epitopes" have been created. A junctional epitope is a potential HLA binding epitope, as predicted, *e.g.*, by motif analysis. Junctional epitopes are generally to be avoided because the recipient may bind to an HLA molecule and generate an immune response to that epitope, which is not present in a native protein sequence. Of particular concern is a junctional epitope that is a "dominant epitope." A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

Peptide epitopes for inclusion in vaccine compositions are, for example, selected from those listed in Tables XXII, XXVI, XXVII, and XXXII. A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response that results in tumor cell killing and reduction of tumor size or mass.

Example 11. Construction of Minigene Multi-Epitope DNA Plasmids

This example provides general guidance for the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of CTL and/or HTL epitopes or epitope analogs as described herein. Expression plasmids have been constructed and evaluated as described, for example, in co-pending U.S.S.N. 09/311,784 filed 5/13/99.

A minigene expression plasmid may include multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes. Preferred epitopes are identified, for example, in Tables XXII, XXVI-XXVIII, and XXXII. HLA class I supermotif or motif-bearing peptide epitopes derived from multiple TAAs are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from multiple tumor antigens to provide broad population coverage, *i.e.* both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

This example illustrates the methods to be used for construction of such a minigene-bearing expression plasmid. Other expression vectors that may be used for minigene compositions are available and known to those of skill in the art.

The minigene DNA plasmid contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by CTL and/or HTL epitopes selected in accordance with principles disclosed herein. The sequence encodes an open reading frame fused to the Myc and His antibody epitope tag coded for by the pcDNA 3.1 Myc-His vector.

Overlapping oligonucleotides, for example eight oligonucleotides, averaging approximately 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides, Kozak sequence, and signal sequence. The final multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95°C for 15 sec, annealing temperature (5° below the lowest calculated Tm of each primer pair) for 30 sec, and 72°C for 1 min.

For the first PCR reaction, 5 μg of each of two oligonucleotides are annealed and extended: Oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 μl reactions containing *Pfu* polymerase buffer (1x= 10 mM KCL, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-chloride, pH 8.75, 2 mM $MgSO_4$, 0.1% Triton X-100, 100 μg/ml BSA), 0.25 mM each dNTP, and 2.5 U of *Pfu* polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product for 25 additional cycles. The full-length product is gel-purified and cloned into pCR-blunt (Invitrogen) and individual clones are screened by sequencing.

Example 12. The plasmid construct and the degree to which it induces immunogenicity.

The degree to which the plasmid construct prepared using the methodology outlined in Example 11 is able to induce immunogenicity is evaluated through *in vivo* injections into mice and subsequent *in vitro* assessment of CTL and HTL activity, which

are analysed using cytotoxicity and proliferation assays, respectively, as detailed *e.g.*, in U.S.S.N. 09/311,784 filed 5/13/99 and Alexander *et al.*, *Immunity* 1:751-761, 1994.

Alternatively, plasmid constructs can be evaluated *in vitro* by testing for epitope presentation by APC following transduction or transfection of the APC with an epitope-expressing nucleic acid construct. Such a study determines "antigenicity" and allows the use of human APC. The assay determines the ability of the epitope to be presented by the APC in a context that is recognized by a T cell by quantifying the density of epitope-HLA class I complexes on the cell surface. Quantitation can be performed by directly measuring the amount of peptide eluted from the APC (*see, e.g.*, Sijts *et al.*, *J. Immunol.* 156:683-692, 1996; Demotz *et al.*, *Nature* 342:682-684, 1989); or the number of peptide-HLA class I complexes can be estimated by measuring the amount of lysis or lymphokine release induced by infected or transfected target cells, and then determining the concentration of peptide necessary to obtained equivalent levels of lysis or lymphokine release (*see, e.g.*, Kageyama *et al.*, *J. Immunol.* 154:567-576, 1995).

To assess the capacity of the pMin minigene construct (*e.g.*, a pMin minigene construct generated as decribed in U.S.S.N. 09/311,784) to induce CTLs *in vivo*, HLA-A11/K^b transgenic mice, for example, are immunized intramuscularly with 100 μg of naked cDNA.. As a means of comparing the level of CTLs induced by cDNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.

Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polyepitopic peptide), then assayed for peptide-specific cytotoxic activity in a ^{51}Cr release assay. The results indicate the magnitude of the CTL response directed against the A3-restricted epitope, thus indicating the *in vivo* immunogenicity of the minigene vaccine and polyepitopic vaccine. It is, therefore, found that the minigene elicits immune responses directed toward the HLA-A3 supermotif peptide epitopes as does the polyepitopic peptide vaccine. A similar analysis is also performed using other HLA-A2 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A2 and HLA-B7 motif or supermotif epitopes.

To assess the capacity of a class II epitope encoding minigene to induce HTLs *in vivo*, I-A^b restricted mice, for example, are immunized intramuscularly with 100 μg of

plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant. CD4+ T cells, *i.e.* HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a ^{3}H-thymidine incorporation proliferation assay, (*see, e.g.*, Alexander et al. Immunity 1:751-761, 1994). The results indicate the magnitude of the HTL response, thus demonstrating the *in vivo* immunogenicity of the minigene.

DNA minigenes, constructed as described in Example 11, may also be evaluated as a vaccine in combination with a boosting agent using a prime boost protocol. The boosting agent may consist of recombinant protein (*e.g.*, Barnett *et al.*, *Aids Res. and Human Retroviruses* 14, *Supplement* 3:S299-S309, 1998) or recombinant vaccinia, for example, expressing a minigene or DNA encoding the complete protein of interest (*see, e.g.*, Hanke *et al.*, *Vaccine* 16:439-445, 1998; Sedegah *et al.*, *Proc. Natl. Acad. Sci USA* 95:7648-53, 1998; Hanke and McMichael, *Immunol. Letters* 66:177-181, 1999; and Robinson *et al.*, *Nature Med.* 5:526-34, 1999).

For example, the efficacy of the DNA minigene may be evaluated in transgenic mice. In this example, A2.1/K^{b} transgenic mice are immunized IM with 100 μg of the DNA minigene encoding the immunogenic peptides. After an incubation period (ranging from 3-9 weeks), the mice are boosted IP with 10^{7} pfu/mouse of a recombinant vaccinia virus expressing the same sequence encoded by the DNA minigene. Control mice are immunized with 100 μg of DNA or recombinant vaccinia without the minigene sequence, or with DNA encoding the minigene, but without the vaccinia boost. After an additional incubation period of two weeks, splenocytes from the mice are immediately assayed for peptide-specific activity in an ELISPOT assay. Additionally, splenocytes are stimulated *in vitro* with the A2-restricted peptide epitopes encoded in the minigene and recombinant vaccinia, then assayed for peptide-specific activity in an IFN-γ ELISA. It is found that the minigene utilized in a prime-boost mode elicits greater immune responses toward the HLA-A2 supermotif peptides than with DNA alone. Such an analysis is also performed using other HLA-A11 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 and HLA-B7 motif or supermotif epitopes.

Example 13. Peptide Composition for Prophylactic Uses

Vaccine compositions of the present invention are used to prevent cancer in persons who are at risk for developing a tumor. For example, a polyepitopic peptide epitope composition (or a nucleic acid comprising the same) containing multiple CTL and HTL epitopes such as those selected in Examples 9 and/or 10, which are also selected to target greater than 80% of the population, is administered to an individual at risk for a cancer, *e.g.*, breast cancer. The composition is provided as a single polypeptide that encompasses multiple epitopes. The vaccine is administered in an aqueous carrier comprised of Freunds Incomplete Adjuvant. The dose of peptide for the initial immunization is from about 1 to about 50,000 μg, generally 100-5,000 μg, for a 70 kg patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient, by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found to be both safe and efficacious as a prophylaxis against cancer.

Alternatively, the polyepitopic peptide composition can be administered as a nucleic acid in accordance with methodologies known in the art and disclosed herein.

Example 14. Polyepitopic Vaccine Compositions Derived from Native TAA Sequences

A native TAA polyprotein sequence is screened, preferably using computer algorithms defined for each class I and/or class II supermotif or motif, to identify "relatively short" regions of the polyprotein that comprise multiple epitopes and is preferably less in length than an entire native antigen. This relatively short sequence that contains multiple distinct, even overlapping, epitopes is selected and used to generate a minigene construct. The construct is engineered to express the peptide, which corresponds to the native protein sequence. The "relatively short" peptide is generally less than 1000, 500, or 250 amino acids in length, often less than 100 amino acids in length, preferably less than 75 amino acids in length, and more preferably less than 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has maximal number of epitopes contained within the sequence, *i.e.*, it has a high concentration of epitopes. As noted herein, epitope motifs may be nested or overlapping (*i.e.,* frame shifted relative to one another). For example, with frame shifted overlapping epitopes, two 9-mer epitopes and one 10-mer epitope can be present in a 10

amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes.

The vaccine composition will preferably include, for example, three CTL epitopes and at least one HTL epitope from TAAs. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent analogs) directs the immune response to multiple peptide sequences that are actually present in native TAAs thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of scale when producing nucleic acid vaccine compositions.

Related to this embodiment, computer programs can be derived in accordance with principles in the art, which identify in a target sequence, the greatest number of epitopes per sequence length.

Example 15. Polyepitopic Vaccine Compositions Directed To Multiple Tumors

The CEA peptide epitopes of the present invention are used in conjunction with peptide epitopes from other target tumor antigens to create a vaccine composition that is useful for the treatment of various types of tumors. For example, a set of TAA epitopes can be selected that allows the targeting of most common epithelial tumors (*see, e.g.*, Kawashima *et al.*, *Hum. Immunol.* 59:1-14, 1998). Such a composition includes epitopes from CEA, HER-2/neu, and MAGE2/3, all of which are expressed to appreciable degrees (20-60%) in frequently found tumors such as lung, breast, and gastrointestinal tumors.

The composition can be provided as a single polypeptide that incorporates the multiple epitopes from the various TAAs, or can be administered as a composition comprising one or more discrete epitopes. Alternatively, the vaccine can be administered as a minigene construct or as dendritic cells which have been loaded with the peptide epitopes *in vitro*.

Targeting multiple tumor antigens is also important to provide coverage of a large fraction of tumors of any particular type. A single TAA is rarely expressed in the majority of tumors of a given type. For example, approximately 50% of breast tumors express CEA, 20% express MAGE3, and 30% express HER-2/neu. Thus, the use of a single antigen for immunotherapy would offer only limited patient coverage. The combination of the three TAAs, however, would address approximately 70% of breast tumors. Furthermore, with the inclusion of CTL epitopes derived from p53, which is overexpressed in approximately 50% of breast tumors, coverage of approximately 85% of all breast tumors could be achieved. A vaccine composition comprising epitopes from multiple tumor antigens also reduces the potential for escape mutants due to loss of expression of an individual tumor antigen.

Example 16. Use of peptides to evaluate an immune response

Peptides of the invention may be used to analyze an immune response for the presence of specific CTL or HTL populations directed to a TAA. Such an analysis may be performed using multimeric complexes as described, *e.g.*, by Ogg *et al., Science* 279:2103-2106, 1998 and Greten *et al., Proc. Natl. Acad. Sci. USA* 95:7568-7573, 1998. In the following example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

In this example, highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") are used for a cross-sectional analysis of, for example, tumor-associated antigen HLA-A*0201-specific CTL frequencies from HLA A*0201-positive individuals at different stages of disease or following immunization using a TAA peptide containing an A*0201 motif. Tetrameric complexes are synthesized as described (Musey *et al., N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A*0201 in this example) and β2-microglobulin are synthesized by means of a prokaryotic expression system. The heavy chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain, β2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Missouri), adenosine 5'triphosphate and magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and

the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300g for 5 minutes and resuspended in 50 μl of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and anti-CD38. The PBMCs are incubated with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A*0201-negative individuals and A*0201-positive uninfected donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that contain epitope-restricted CTLs, thereby readily indicating the extent of immune response to the TAA epitope, and thus the stage of tumor progression or exposure to a vaccine that elicits a protective or therapeutic response.

Example 17. Use of Peptide Epitopes to Evaluate Recall Responses

The peptide epitopes of the invention are used as reagents to evaluate T cell responses, such as acute or recall responses, in patients. Such an analysis may be performed on patients who are in remission, have a tumor, or who have been vaccinated with a TAA vaccine.

For example, the class I restricted CTL response of persons who have been vaccinated may be analyzed. The vaccine may be any TAA vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide epitopes of the invention that, optimally, bear supermotifs to provide cross-reactivity with multiple HLA supertype family members, are then used for analysis of samples derived from individuals who bear that HLA type.

PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, MO), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2mM), penicillin (50U/ml), streptomycin (50 μg/ml), and Hepes (10mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using microculture formats. A synthetic peptide comprising an epitope of the invention is added at 10 μg/ml to each well and HBV core 128-140 epitope is added at 1 μg/ml to each well as a source of T cell help during the first week of stimulation.

In the microculture format, 4 x 10^5 PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 µl/well of complete RPMI. On days 3 and 10, 100 µl of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and 10^5 irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response requires two or more of the eight replicate cultures to display greater than 10% specific ^{51}Cr release, based on comparison with uninfected control subjects as previously described (Rehermann, *et al., Nature Med.* 2:1104,1108, 1996; Rehermann *et al., J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann *et al. J. Clin. Invest.* 98:1432-1440, 1996).

Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, MA) or established from the pool of patients as described (Guilhot, *et al. J. Virol. 66*:2670-2678, 1992).

Cytotoxicity assays are performed in the following manner. Target cells consist of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with the synthetic peptide epitope of the invention at 10 µM, and labeled with 100 µCi of ^{51}Cr (Amersham Corp., Arlington Heights, IL) for 1 hour after which they are washed four times with HBSS.

Cytolytic activity is determined in a standard 4 hour, split-well ^{51}Cr release assay using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at effector/target (E/T) ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula: 100 x [(experimental release-spontaneous release)/maximum release-spontaneous release)]. Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, MO). Spontaneous release is <25% of maximum release for all experiments.

The results of such an analysis indicate the extent to which HLA-restricted CTL populations have been stimulated by previous exposure to the TAA or TAA vaccine.

The class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of 1.5x10^5 cells/well and are stimulated with 10 µg/ml synthetic peptide, whole antigen, or PHA. Cells are routinely plated in replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10U/ml IL-2. Two days later, 1 µCi

^{3}H-thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for ^{3}H-thymidine incorporation. Antigen-specific T cell proliferation is calculated as the ratio of ^{3}H-thymidine incorporation in the presence of antigen divided by the ^{3}H-thymidine incorporation in the absence of antigen.

Example 18. Induction Of Specific CTL Response In Humans

A human clinical trial for an immunogenic composition comprising CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study. Such a trial is designed, for example, as follows:

A total of about 27 subjects are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 μg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 μg peptide composition;

Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 μg of peptide composition.

After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage. Additional booster inoculations can be administered on the same schedule.

The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

The vaccine is found to be both safe and efficacious.

Example 19. Therapeutic Use in Cancer Patients

Evaluation of vaccine compositions are performed to validate the efficacy of the CTL-HTL peptide compositions in cancer patients. The main objectives of the trials are to determine an effective dose and regimen for inducing CTLs in cancer patients, to establish the safety of inducing a CTL and HTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of cancer patients, as manifested by a reduction in tumor cell numbers. Such a study is designed, for example, as follows:

The studies are performed in multiple centers. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster shot of the same dose. The dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects (severity and reversibility) are recorded.

There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65, include both males and females (unless the tumor is sex-specific, *e.g.*, breast or prostate cancer), and represent diverse ethnic backgrounds.

Example 20. Induction of CTL Responses Using a Prime Boost Protocol

A prime boost protocol similar in its underlying principle to that used to evaluate the efficacy of a DNA vaccine in transgenic mice, which was described in Example 12, may also be used for the administration of the vaccine to humans. Such a vaccine regimen may include an initial administration of, for example, naked DNA followed by a boost using recombinant virus encoding the vaccine, or recombinant protein/polypeptide or a peptide mixture administered in an adjuvant.

For example, the initial immunization may be performed using an expression vector, such as that constructed in Example 11, in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 μg) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is then administered. The booster can be recombinant fowlpox virus administered at a dose of $5\text{-}10^7$ to $5x10^9$ pfu. An alternative recombinant virus, such as an MVA, canarypox, adenovirus, or adeno-associated virus, can also be used for the booster, or the polyepitopic protein or a mixture of the peptides can be

administered. For evaluation of vaccine efficacy, patient blood samples will be obtained before immunization as well as at intervals following administration of the initial vaccine and booster doses of the vaccine. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

Analysis of the results will indicate that a magnitude of response sufficient to achieve protective immunity against cancer is generated.

Example 21. Administration of Vaccine Compositions Using Dendritic Cells

Vaccines comprising peptide epitopes of the invention may be administered using dendritic cells. In this example, the immunogenic peptide epitopes are used to elicit a CTL and/or HTL response *ex vivo*.

Ex vivo CTL or HTL responses to a particular tumor-associated antigen are induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptides. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cells, *i.e.*, tumor cells.

Alternatively, the peptide-pulsed dendritic cells may be administered to the patient to stimulate a CTL response *in vivo*. In this method, dendritic cells are isolated as described in Example 3. The dendritic cell population is expanded and pulsed with a vaccine comprising peptide CTL and HTL epitopes of the invention. The dendritic cells are infused back into the patient to elicit CTL and HTL responses *in vivo*. The induced CTL and HTL then destroy (CTL) or facilitate destruction (HTL) of the specific target tumor cells that bear the proteins from which the epitopes in the vaccine are derived.

Example 22. Alternative Method of Identifying Motif-Bearing Peptides

Another way of identifying motif-bearing peptides is to elute them from cells bearing defined MHC molecules. For example, EBV transformed B cell lines used for tissue typing, have been extensively characterized to determine which HLA molecules they express. In certain cases these cells express only a single type of HLA molecule. These cells can then be infected with a pathogenic organism or transfected with nucleic

acids that express the tumor antigen of interest. Thereafter, peptides produced by endogenous antigen processing of peptides produced consequent to infection (or as a result of transfection) will bind to HLA molecules within the cell and be transported and displayed on the cell surface.

The peptides are then eluted from the HLA molecules by exposure to mild acid conditions and their amino acid sequence determined, e.g., by mass spectral analysis (*e.g.*, Kubo *et al.*, *J. Immunol.* 152:3913, 1994). Because, as disclosed herein, the majority of peptides that bind a particular HLA molecule are motif-bearing, this is an alternative modality for obtaining the motif-bearing peptides correlated with the particular HLA molecule expressed on the cell.

Alternatively, cell lines that do not express any endogenous HLA molecules can be transfected with an expression construct encoding a single HLA allele. These cells may then be used as described, *i.e.*, they may be infected with a pathogenic organism or transfected with nucleic acid encoding an antigen of interest to isolate peptides corresponding to the pathogen or antigen of interest that have been presented on the cell surface. Peptides obtained from such an analysis will bear motif(s) that correspond to binding to the single HLA allele that is expressed in the cell.

As appreciated by one in the art, one can perform a similar analysis on a cell bearing more than one HLA allele and subsequently determine peptides specific for each HLA allele expressed. Moreover, one of skill would also recognize that means other than infection or transfection, such as loading with a protein antigen, can be used to provide a source of antigen to the cell.

The above examples are provided to illustrate the invention but not to limit its scope. For example, the human terminology for the Major Histocompatibility Complex, namely HLA, is used throughout this document. It is to be appreciated that these principles can be extended to other species as well. Thus, other variants of the invention will be readily apparent to one of ordinary skill in the art and are encompassed by the appended claims. All publications, patents, and patent application cited herein are hereby incorporated by reference for all purposes.

TABLE I

SUPERMOTIFS	POSITION 2 (Primary Anchor)	POSITION 3 (Primary Anchor)	POSITION C Terminus (Primary Anchor)
A1	**TI***LVMS*		**FWY**
A2	**LIVM***ATQ*		**IV***MATL*
A3	**VSMA***TLI*		**RK**
A24	**YF***WIVLMT*		**FI***YWLM*
B7	**P**		**VILF***MWYA*
B27	**RHK**		**FYL***WMIVA*
B44	**E***D*		**FWYLIMVA**
B58	**ATS**		**FWY***LIVMA*
B62	**QL***IVMP*		**FWY***MIVLA*
MOTIFS			
A1	**TSM**		**Y**
A1		**DE***AS*	**Y**
A2.1	**LM***VQIAT*		**V***LIMAT*
A3	**LMVISATF***CGD*		**KYR***HFA*
A11	**VTMLISAGN***CDF*		**K***RYH*
A24	**YFW***M*		**FLIW**
A*3101	**MVT***ALIS*		**R***K*
A*3301	**MVALF***IST*		**RK**
A*6801	**AVT***MSLI*		**RK**
B*0702	**P**		**LMF***WYAIV*
B*3501	**P**		**LMFWY***IVA*
B51	**P**		**LIVF***WYAM*
B*5301	**P**		**IMFWY***ALV*
B*5401	**P**		**ATIV***LMFWY*

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE Ia

SUPERMOTIFS	POSITION 2 (Primary Anchor)	POSITION 3 (Primary Anchor)	POSITION C Terminus (Primary Anchor)
A1	**TI***LVMS*		**FWY**
A2	*VQAT*		**V***LIMAT*
A3	**VSMA***TLI*		**RK**
A24	**YF***WIVLMT*		**FI***YWLM*
B7	**P**		**VILF***MWYA*
B27	**RHK**		**FYL***WMIVA*
B58	**ATS**		**FWY***LIVMA*
B62	**QL***IVMP*		**FWY***MIVLA*
MOTIFS			
A1	**TSM**		**Y**
A1		**DE***AS*	**Y**
A2.1	*VQAT**		**V***LIMAT*
A3.2	**LMVISATF***CGD*		**KYR***HFA*
A11	**VTMLISAGN***CDF*		**K***RHY*
A24	**YFW**		**FLIW**

*If 2 is V, or Q, the C-term is not L

Bolded residues are preferred, italicized residues are less preferred: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE II

		POSITION								
		1	2	3	4	5	6	7	8	C-terminus
SUPERMOTIFS										
A1			1° Anchor TI*LVMS*							1° Anchor FWY
A2			1° Anchor LIVM*ATQ*							1° Anchor LIVMAT
A3	preferred		1° Anchor VSMA*TLI*	YFW (4/5)			YFW (3/5)	YFW (4/5)	P (4/5)	1°Anchor RK
	deleterious	DE (3/5); P (5/5)		DE (4/5)						
A24			1° Anchor YF*WIVLM* *T*							1° Anchor FI*YWLM*
B7	preferred	FWY (5/5) LIVM (3/5)	1°Anchor P	FWY (4/5)					FWY (3/5)	1°Anchor VILF*MWYA*
	deleterious	DE (3/5); P(5/5); G(4/5); A(3/5); QN (3/5)				DE (3/5)	G (4/5)	QN (4/5)	DE (4/5)	
B27			1° Anchor RHK							1° Anchor FYL*WMIVA*
B44			1° Anchor E*D*							1° Anchor FWYLIMVA
B58			1° Anchor ATS							1° Anchor FWY*LIVMA*
B62			1° Anchor QL*IVMP*							1° Anchor FWY*MIVLA*

		POSITION								
		1	2	3	4	5	6	7	8	C-terminus

		POSITION								
		1	2	3	4	5	6	7	8	C-terminus
MOTIFS										
A1 9-mer	preferred	GFYW	1°Anchor STM	DEA	YFW		P	DEQN	YFW	1°Anchor Y
	deleterious	DE		RHKLIVM P	A	G	A			
A1 9-mer	preferred	GRHK	ASTCLIV M	1°Anchor DEAS	GSTC		ASTC	LIVM	DE	1°Anchor Y
	deleterious	A	RHKDEPY FW		DE	PQN	RHK	PG	GP	

		POSITION									
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A1 10-mer	peferred	YFW	1°Anchor STM	DEAQN	A	YFWQN		PASTC	GDE	P	1°Anchor Y
	deleterious	GP		RHKGLIVM	DE	RHK	QNA	RHKYFW	RHK	A	
A1 10-mer	preferred	YFW	STCLIVM	1°Anchor DE*AS*	A	YFW		PG	G	YFW	1°Anchor Y
	deleterious	RHK	RHKDEPYFW			P	G		PRHK	QN	
A2.1 9-mer	preferred	YFW	1°Anchor LM*IVQAT*	YFW	STC	YFW		A	P	1°Anchor V*LIMAT*	
	deleterious	DEP		DERKH			RKH	DERKH			
A2.1 10-mer	preferred	AYFW	1°Anchor LM*IVQAT*	LVIM	G		G		FYWL VIM		1°Anchor V*LIMAT*
	deleterious	DEP		DE	RKHA	P		RKH	DERKH	RKH	

		POSITION									
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A3	preferred	RHK	1°Anchor LMVISAT FCGD	YFW	PRHKYFW	A	YFW		P	1°Anchor KYRHFA	
	deleterious	DEP		DE							
A11	preferred	A	1°Anchor VTLMISA GNCDF	YFW	YFW	A	YFW	YFW	P	1°Anchor KRYH	
	deleterious	DEP						A	G		
A24 9-mer	preferred	YFWRHK	1°Anchor YFWM		STC			YFW	YFW	1°Anchor FLIW	
	deleterious	DEG		DE	G	QNP	DERHK	G	AQN		
A24 10-mer	preferred		1°Anchor YFWM		P	YFWP		P			1°Anchor FLIW
	deleterious			GDE	QN	RHK	DE	A	QN	DEA	

		POSITION									
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A3101	preferred	RHK	1°Anchor MVT*ALIS*	YFW	P		YFW	YFW	AP	1°Anchor R*K*	
	deleterious	DEP		DE		ADE	DE	DE	DE		
A3301	preferred		1°Anchor MVALF*IS T*	YFW				AYFW		1°Anchor RK	
	deleterious	GP		DE							
A6801	preferred	YFWSTC	1°Anchor AVT*MSLI*			YFWLIV M		YFW	P	1°Anchor RK	
	deleterious	GP		DEG		RHK			A		
B0702	preferred	RHKFWY	1°Anchor P	RHK		RHK	RHK	RHK	PA	1°Anchor LMF*WYAIV*	
	deleterious	DEQNP		DEP	DE	DE	GDE	QN	DE		
B3501	preferred	FWYLIVM	1°Anchor P	FWY				FWY		1°Anchor LMFWY*IVA*	
	deleterious	AGP				G	G				

		POSITION									
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
B51	preferred	LIVMFWY	1°Anchor P	FWY	STC	FWY		G	FWY	1°Anchor LIVF*WYAM*	
	deleterious	AGPDERHKSTC				DE	G	DEQN	GDE		
B5301	preferred	LIVMFWY	1°Anchor P	FWY	STC	FWY		LIVMFWY	FWY	1°Anchor IMFWY*ALV*	
	deleterious	AGPQN					G	RHKQN	DE		
B5401	preferred	FWY	1°Anchor P	FWYLIVM		LIVM		ALIVM	FWYAP	1°Anchor ATIV*LMFW Y*	
	deleterious	GPQNDE		GDESTC		RHKDE	DE	QNDGE	DE		

Italicized residues indicate less preferred or "tolerated" residues.
The information in Table II is specific for 9-mers unless otherwise specified.

TABLE III

MOTIFS		1° anchor 1	2	3	4	5	1° anchor 6	7	8	9
						POSITION				
DR4	preferred	FMY*LIVW*	M	T		I	VST*CPALIM*	MH		MH
	deleterious				W			R		WDE
DR1	preferred	MF*LIVWY*			PAMQ		VMAT*SPLIC*	M		AVM
	deleterious		C	CH	FD	CWD		GDE	D	
DR7	preferred	MF*LIVWY*	M	W	A		IVMSA*CTPL*	M		IV
	deleterious		C		G			GRD	N	G
DR Supermotif		MF*LIVWY*					VMSTA*CPLI*			

DR3 MOTIFS	1° anchor 1	2	3	1° anchor 4	5	1° anchor 6
motif a preferred	LIVMFY			D		
motif b preferred	LIVMFAY			DNQEST		KRH

Italicized residues indicate less preferred or "tolerated" residues.

Table IV. HLA Class I Standard Peptide Binding Affinity.

ALLELE	STANDARD PEPTIDE	SEQUENCE	STANDARD BINDING AFFINITY (nM)
A*0101	944.02	YLEPAIAKY	25
A*0201	941.01	FLPSDYFPSV	5.0
A*0202	941.01	FLPSDYFPSV	4.3
A*0203	941.01	FLPSDYFPSV	10
A*0205	941.01	FLPSDYFPSV	4.3
A*0206	941.01	FLPSDYFPSV	3.7
A*0207	941.01	FLPSDYFPSV	23
A*6802	1072.34	YVIKVSARV	8.0
A*0301	941.12	KVFPYALINK	11
A*1101	940.06	AVDLYHFLK	6.0
A*3101	941.12	KVFPYALINK	18
A*3301	1083.02	STLPETYVVRR	29
A*6801	941.12	KVFPYALINK	8.0
A*2402	979.02	AYIDNYNKF	12
B*0702	1075.23	APRTLVYLL	5.5
B*3501	1021.05	FPFKYAAAF	7.2
B51	1021.05	FPFKYAAAF	5.5
B*5301	1021.05	FPFKYAAAF	9.3
B*5401	1021.05	FPFKYAAAF	10

Table V. HLA Class II Standard Peptide Binding Affinity.

Allele	Nomenclature	Standard Peptide	Sequence	Binding Affinity (nM)
DRB1*0101	DR1	515.01	PKYVKQNTLKLAT	5.0
DRB1*0301	DR3	829.02	YKTIAFDEEARR	300
DRB1*0401	DR4w4	515.01	PKYVKQNTLKLAT	45
DRB1*0404	DR4w14	717.01	YARFQSQTTLKQKT	50
DRB1*0405	DR4w15	717.01	YARFQSQTTLKQKT	38
DRB1*0701	DR7	553.01	QYIKANSKFIGITE	25
DRB1*0802	DR8w2	553.01	QYIKANSKFIGITE	49
DRB1*0803	DR8w3	553.01	QYIKANSKFIGITE	1600
DRB1*0901	DR9	553.01	QYIKANSKFIGITE	75
DRB1*1101	DR5w11	553.01	QYIKANSKFIGITE	20
DRB1*1201	DR5w12	1200.05	EALIHQLKINPYVLS	298
DRB1*1302	DR6w19	650.22	QYIKANAKFIGITE	3.5
DRB1*1501	DR2w2β1	507.02	GRTQDENPVVHFFKNIV TPRTPPP	9.1
DRB3*0101	DR52a	511	NGQIGNDPNRDIL	470
DRB4*0101	DRw53	717.01	YARFQSQTTLKQKT	58
DRB5*0101	DR2w2β2	553.01	QYIKANSKFIGITE	20

The "Nomenclature" column lists the allelic designations used in Tables XIX and XX.

Table VI

HLA-supertype	Allelle-specific HLA-supertype members	
	Verified[a]	Predicted[b]
A1	A*0101, A*2501, A*2601, A*2602, A*3201	A*0102, A*2604, A*3601, A*4301, A*8001
A2	A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, A*6901	A*0208, A*0210, A*0211, A*0212, A*0213
A3	A*0301, A*1101, A*3101, A*3301, A*6801	A*0302, A*1102, A*2603, A*3302, A*3303, A*3401, A*3402, A*6601, A*6602, A*7401
A24	A*2301, A*2402, A*3001	A*2403, A*2404, A*3002, A*3003
B7	B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, B*7801	B*1511, B*4201, B*5901
B27	B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, B*7301	B*2701, B*2707, B*2708, B*3802, B*3903, B*3904, B*3905, B*4801, B*4802, B*1510, B*1518, B*1503
B44	B*1801, B*1802, B*3701, B*4402, B*4403, B*4404, B*4001, B*4002, B*4006	B*4101, B*4501, B*4701, B*4901, B*5001
B58	B*5701, B*5702, B*5801, B*5802, B*1516, B*1517	
B62	B*1501, B*1502, B*1513, B*5201	B*1301, B*1302, B*1504, B*1505, B*1506, B*1507, B*1515, B*1520, B*1521, B*1512, B*1514, B*1510

a. Verified alleles inclueds alleles whose specificity has been determined by pool sequencing analysis, peptide binding assays, or by analysis of the sequences of CTL epitopes.

b. Predicted alleles are alleles whose specificity is predicted on the basis of B and F pocket structure to overlap with the supertype specificity.

Table VII

CEA A01 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0101	SEQ ID NO.
ASNPPAQY	440	8	0.0120	1
ASNPPAQYSW	440	10		2
ASNPPAQYSWF	262	11		3
ASNPSPQY	618	8	0.0085	4
ASNPSPQYSW	618	10		5
ATGQFRVY	134	8	-0.0021	6
DLVNEEATGQF	128	11		7
DSVILNVLY	227	9	-0.0021	8
EIQNTTYLW	348	9		9
EIQNTTYLWW	348	10		10
ESPSAPPHRW	2	10		11
ETQDATYLW	170	9		12
ETQDATYLWW	170	10		13
GIPQQHTQVLF	631	11		14
GTFQQSTQELF	275	11		15
GTQQATPGPAY	85	11	0.0069	16
HLFGYSWY	61	8		17
HSASNPSPQY	616	10	0.3400	18
HSDPVILNVLY	403	11	0.9700	19
HQNDTGF	112	8		20
HQNDTGFY	112	9		21
HSPPDSSY	597	9	0.0021	22
ISPLNTSY	242	8	-0.0021	23
ISPPDSSY	598	8	-0.0021	24
ISPSYTYY	420	8	0.0030	25
ITEKNSGLY	467	9	0.0390	26
ITPNNNGTY	645	9	0.0049	27
ITVNNSGSY	289	9	0.0100	28
ITVYAEPPKPF	316	11		29
KITPNNNGTY	644	10		30
KLTIESTPF	35	9		31
LLLTASLLTF	18	10		32
LLLTASLLTFW	18	11		33
LLTASLLTF	19	9		34
LLTASLLTFW	19	10		35
LLVHNLPQHLF	53	11		36
LSNGNRTLTLF	549	11		37
LSVTRNDVGPY	381	11	0.0100	38
LTASLLTF	20	8		39
LTASLLTFW	20	9		40
LTIESTPF	36	8		41
LVHNLPQHLF	54	10		42
LVNEEATGQF	129	10		43
NHQNDTGF	111	9		44
NHQNDTGFY	111	10		45
NIQQHTQELF	454	10		46
NITEKNSGLY	466	10		47
NITVNNSGSY	288	10		48
NLPQHLFGY	57	9		49
NLPQHLFGYSW	57	11		50

Table VII
CEA A01 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0101	SEQ ID NO.
NVTRNDARAY	560	10		51
NVTRNDTASY	204	10		52
PHSPPDSSY	596	10		53
PSAPPHRW	4	8		54
PTISPLNTSY	240	10	0.0250	55
PTISPSYTY	418	9	0.0035	56
PTISPSYTYY	418	10	0.0770	57
PVEDKDAVAF	512	10		58
PVILNVLY	406	8		59
PVTLDVLY	584	8		60
RLLLTASLLTF	17	11		61
RSDPVTLDVLY	581	11	3.2000	62
RSDSVILNVLY	225	11	0.5300	63
RTTVTTITVY	310	10	0.0041	64
RVDGNRQHGY	72	11	0.0850	65
SVILNVLY	228	8		66
SVTRNDVGPY	382	10		67
TISPLNTSY	241	9	0.0024	68
TISPSYTY	419	8	0.0038	69
TISPSYTYY	419	9	0.0240	70
TTVTTITVY	311	9	0.0011	71
TVNNSGSY	290	8		72
TVTTITVY	312	8		73
TVYAEPPKPF	317	10		74
VTRNDARAY	561	9	0.0011	75
VTRNDTASY	205	9	0.0011	76
VTRNDVGPY	383	9	-0.0021	77
YSGREHY	95	8	0.0150	78
YSWFVNGTF	269	9		79

Table VIII

CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
ALTCEPEI	342	8	0.0002					80
ALTCEPEIQNT	342	11	-0.0001					81
AQNTTYLWWV	527	10						82
AQYSWFVNGT	267	10						83
AQYSWLIDGNI	445	11						84
ATGQFRVYPEL	134	11	-0.0001					85
ATGRNNSI	661	8	-0.0002					86
ATGRNNSIV	661	9	-0.0002					87
ATVGIMIGV	687	9	0.0280	0.1100	0.1300	0.1500	1.6000	88
ATVGIMIGVL	687	10	0.0007					89
ATVGIMIGVLV	687	11	0.0160					90
AVAFTCEPEA	518	10	0.0003					91
AVAFTCEPET	162	10						92
AVALTCEPEI	340	10	0.0002					93
CIPWQRLL	12	8	-0.0002					94
CIPWQRLLL	12	9	0.0002					95
CIPWQRLLLT	12	10	0.0031					96
CIPWQRLLLTA	12	11	0.0003					97
CQAHNSDT	299	8						98
CQAHNSDTGL	299	10						99
DAPTISPL	238	8	-0.0002					100
DAPTISPLNT	238	10	-0.0002					101
DARAYVCGI	565	9	-0.0002					102
DATYLWWV	173	8	0.0001					103
DAVAFTCEPEA	517	11	-0.0001					104
DAVAFTCEPET	161	11						105
DAVALTCEPEI	339	11	-0.0001					106
DLVNEEAT	128	8						107
DTASYKCET	209	9	-0.0002					108
DTGFYTLHV	116	9	0.0009					109
DTGFYTLHVI	116	10	-0.0002					110
DTGLNRTT	305	8	-0.0002					111
DTGLNRTTV	305	9	-0.0002					112
DTGLNRTTVT	305	10	-0.0002					113
DTGLNRTTVTT	305	11	0.0001					114
DVGPYECGI	387	9	-0.0002					115
DVLYGPDT	588	8	-0.0002					116
DVLYGPDTPI	588	10	0.0003					117
DVLYGPDTPII	588	11	0.0001					118
EAQNTTYL	526	8	-0.0002					119
EAQNTTYLWWV	526	11	0.0011					120
EATGQFRV	133	8	0.0001					121
EIIYPNASL	99	9	-0.0002					122
EIIYPNASLL	99	10	-0.0002					123
EIIYPNASLLI	99	11	0.0004					124
EIQNTTYL	348	8	-0.0002					125
EIQNTTYLWWV	348	11	0.0004					126
ELFIPNIT	283	8						127
ELFIPNITV	283	9						128
ELFISNIT	461	8	-0.0002					129

Table VIII

CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
ELSVDHSDPV	398	10	0.0001					130
ELSVDHSDPVI	398	11	-0.0001					131
ETQDATYL	170	8	-0.0002					132
ETQDATYLWWV	170	11	0.0002					133
ETQNPVSA	216	8	-0.0002					134
EVLLLVHNL	50	9						135
FITSNNSNPV	326	10	0.0001					136
FQQSTQEL	277	8						137
FQQSTQELFI	277	10						138
FTCEPEAQNT	521	10	0.0003					139
FTCEPEAQNTT	521	11	0.0059					140
FTCEPETQDA	165	10	-0.0002					141
FTCEPETQDAT	165	11	0.0005					142
FVNGTFQQST	272	10	0.0003					143
GANLNLSCHSA	608	11	-0.0001					144
GATVGIMI	686	8	-0.0002					145
GATVGIMIGV	686	10	0.0006					146
GATVGIMIGVL	686	11	0.0051					147
GIMIGVLV	690	8	0.0089					148
GIMIGVLVGV	690	10	0.0880	0.0110	0.1500	0.0250	0.0260	149
GIMIGVLVGVA	690	11	0.0015					150
GIPQQHTQV	631	9	0.0002					151
GIPQQHTQVL	631	10	-0.0002					152
GIQNELSV	394	8	0.0001					153
GIQNSVSA	572	8	-0.0002					154
GLNRTTVT	307	8						155
GLNRTTVTT	307	9	0.0011					156
GLNRTTVTTI	307	10	0.0004					157
GLNRTTVTTIT	307	11	0.0001					158
GLSAGATV	682	8	0.0008					159
GLSAGATVGI	682	10	0.0037					160
GLSAGATVGIM	682	11	0.0001					161
GLYTCQANNSA	473	11	0.0290					162
GQFRVYPEL	136	9						163
GQSLPVSPRL	538	10						164
GTFQQSTQEL	275	10						165
GTQQATPGPA	85	10						166
GTSPGLSA	678	8						167
GTSPGLSAGA	678	10	-0.0002					168
GTSPGLSAGAT	678	11	-0.0001					169
GTYACFVSNL	651	10	0.0002					170
GTYACFVSNLA	651	11	0.0004					171
GVLVGVAL	694	8	-0.0002					172
GVLVGVALI	694	9	0.0030					173
GVNLSLSCHA	430	10	-0.0001					174
GVNLSLSCHAA	430	11	0.0022					175
HAASNPPA	438	8						176
HTQELFISNI	458	10	-0.0001					177
HTQELFISNIT	458	11	0.0013					178
HTQVLFIA	636	8	0.0036					179

TableVIII
CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
IITQVLFIAKI	636	10	0.0012					180
IITQVLFIAKIT	636	11	0.0059					181
IIVIKSDLV	123	8	-0.0002					182
IAKITPNNNGT	642	11	-0.0001					183
IIGYVIGT	79	8	0.0005					184
IIGYVIGTQQA	79	11	-0.0001					185
IIQNDTGFYT	112	10	0.0011					186
IIQNDTGFYTL	112	11	0.0130					187
IISPPDSSYL	597	10	0.0003					188
IIYPNASL	100	8	-0.0002					189
IIYPNASLL	100	9	0.0034					190
IIYPNASLLI	100	10	0.0058					191
ILNVLYGPDA	230	10	0.0007					192
IMIGVLVGV	691	9	0.1500					193
IMIGVLVGVA	691	10	0.0160					194
IMIGVLVGVAL	691	11	0.0029					195
IQNDTGFYT	113	9						196
IQNDTGFYTL	113	10						197
IQNIIQNDT	109	9						198
IQNTTYLWWV	349	10						199
IQQIITQEL	455	8						200
IQQIITQELFI	455	10						201
ITEKNSGL	467	8	-0.0002					202
ITEKNSGLYT	467	10	-0.0002					203
ITPNNNGT	645	8	-0.0002					204
ITPNNNGTYA	645	10	0.0002					205
ITSNNSNPV	327	9	0.0006					206
ITVNNSGSYT	289	10						207
ITVSASGT	672	8	-0.0002					208
IVKSITVSA	668	9	-0.0002					209
KITPNNNGT	644	9	-0.0002					210
KITPNNNGTYA	644	11	0.0002					211
KLTIESTPFNV	35	11						212
KTITVSAEL	492	9	0.0020					213
LATGRNNSI	660	9	-0.0002					214
LATGRNNSIV	660	10	-0.0002					215
LIDGNIQQIIT	450	10	-0.0002					216
LIQNIIQNDT	108	10	0.0003					217
LLIQNIIQNDT	107	11	0.0140					218
LLLTASLL	18	8						219
LLLTASLLT	18	9						220
LLLVIINLPQIIL	52	11	0.0011					221
LLSVTRNDV	380	9	0.0003					222
LLTASLLT	19	8						223
LLTFWNPPT	24	9	0.0260					224
LLTFWNPPTT	24	10						225
LLTFWNPPTTA	24	11						226
LLVIINLPQIIL	53	10	0.0008					227
LQLSNDNRT	369	9						228
LQLSNDNRTL	369	10						229

Table VIII
CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
LQLSNDNRTLT	369	11						230
LQLSNGNRT	547	9						231
LQLSNGNRTL	547	10						232
LQLSNGNRTLT	547	11						233
LTCEPEIQNT	343	10	-0.0002					234
LTCEPEIQNTT	343	11	-0.0001					235
LTFWNPPT	25	8						236
LTFWNPPTT	25	9						237
LTFWNPPTTA	25	10						238
LTIESTPFNV	36	10						239
LTIESTPFNVA	36	11						240
LTLFNVTRNDA	556	11	0.0004					241
LTLFNVTRNDT	200	11	-0.0001					242
LTLLSVTRNDV	378	11	0.0150					243
LVIINLPQIL	54	9	-0.0002					244
MIGVLVGV	692	8	0.0120					245
MIGVLVGVA	692	9	0.0009					246
MIGVLVGVAL	692	10	0.0004					247
MIGVLVGVALI	692	11	0.0025					248
NASLLIQNI	104	9	-0.0002					249
NASLLIQNII	104	10	-0.0002					250
NIIQNDTGFYT	111	11	0.0006					251
NIQQIITQEL	454	9	0.0002					252
NIQQHTQELFI	454	11	0.0001					253
NITEKNSGL	466	9	-0.0002					254
NITEKNSGLYT	466	11	-0.0001					255
NITVNNSGSYT	288	11						256
NLATGRNNSI	659	10	-0.0002					257
NLATGRNNSIV	659	11	0.0001					258
NLNLSCHA	254	8						259
NLNLSCHAA	254	9						260
NLNLSCHSA	610	9	0.0003					261
NLSLSCHA	432	8	-0.0002					262
NLSLSCHAA	432	9	0.0110	0.0015	0.0069	0.0002	0.0003	263
NQSLPVSPRL	360	10						264
NTSYRSGENL	246	10	-0.0002					265
NTTYLWWV	529	8						266
NVAEGKEV	44	8						267
NVAEGKEVL	44	9						268
NVAEGKEVLL	44	10						269
NVAEGKEVLLL	44	11						270
NVLYGPDA	232	8	0.0001					271
NVLYGPDAPT	232	10	-0.0002					272
NVLYGPDAPTI	232	11	0.0001					273
NVLYGPDDPT	410	10	-0.0002					274
NVLYGPDDPTI	410	11	0.0013					275
NVTRNDARA	560	9	-0.0002					276
NVTRNDARAYV	560	11	-0.0001					277
NVTRNDTA	204	8	-0.0002					278
PAQYSWFV	266	8	-0.0002					279

Table VIII

CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
PAQYSWFVNGT	266	11	0.0007					280
PAQYSWLI	444	8						281
PAYSGREI	93	8	-0.0002					282
PAYSGREII	93	9	-0.0002					283
PHSPPDSSYL	596	11	-0.0001					284
PQQHTQVL	633	8						285
PQQHTQVLFI	633	10						286
PQQHTQVLFIA	633	11						287
PQYSWRINGI	623	10						288
PTISPLNT	240	8	-0.0002					289
PTISPSYT	418	8	-0.0002					290
PTTAKLTI	31	8						291
PTTAKLTIEST	31	11						292
PVEDEDAV	334	8	0.0002					293
PVEDEDAVA	334	9	-0.0002					294
PVEDEDAVAL	334	10	-0.0002					295
PVEDEDAVALT	334	11	-0.0001					296
PVEDKDAV	512	8						297
PVEDKDAVA	512	9						298
PVEDKDAVAFT	512	11						299
PVSARRSDSV	220	10	-0.0002					300
PVSARRSDSVI	220	11	-0.0001					301
PVSPRLQL	542	8						302
QAHNSDTGL	300	9	-0.0002					303
QIIGYVIGT	78	9	0.0270	0.0780	0.0730	0.1200	0.2600	304
QLSNDNRT	370	8	-0.0002					305
QLSNDNRTL	370	9	0.0001					306
QLSNDNRTLT	370	10	-0.0002					307
QLSNDNRTLTL	370	11	0.0001					308
QLSNGNRT	548	8						309
QLSNGNRTL	548	9						310
QLSNGNRTLT	548	10						311
QLSNGNRTLTL	548	11						312
QQATPGPA	87	8						313
QQHTQELFI	456	9						314
QQHTQVLFI	634	9						315
QQHTQVLFIA	634	10						316
QQSTQELFI	278	9						317
QVLFIAKI	638	8	0.0007					318
QVLFIAKIT	638	9	0.0008					319
RAYVCGIQNSV	567	11	0.0099					320
RINGIPQQHT	628	10	-0.0002					321
RLLLTASL	17	8	0.0023					322
RLLLTASLL	17	9	0.0068					323
RLLLTASLLT	17	10	0.0036					324
RLQLSNDNRT	368	10	-0.0002					325
RLQLSNDNRTL	368	11	0.0001					326
RLQLSNGNRT	546	10						327
RLQLSNGNRTL	546	11						328
RQIIGYVI	77	8						329

Table VIII

CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
RQIIGYVIGT	77	10						330
RTLTLFNV	554	8	0.0078					331
RTLTLFNVT	554	9	-0.0002					332
RTLTLLSV	376	8						333
RTLTLLSVT	376	9						334
RTTVKTIT	488	8	-0.0002					335
RTTVKTITV	488	9	-0.0002					336
RTTVKTITVSA	488	11	0.0064					337
RTTVTTIT	310	8	-0.0002					338
RTTVTTITV	310	9	0.0012					339
RTTVTTITVYA	310	11	0.0020					340
RVDGNRQI	72	8						341
RVDGNRQII	72	9	-0.0002					342
RVYPELPKPSI	139	11	-0.0001					343
SAELPKPSI	497	9	-0.0002					344
SAGATVGI	684	8	-0.0002					345
SAGATVGIM	684	9	-0.0002					346
SAGATVGIMI	684	10	-0.0002					347
SANRSDPV	578	8						348
SANRSDPVT	578	9						349
SANRSDPVTL	578	10						350
SAPPHRWCI	5	9	-0.0002					351
SARRSDSV	222	8	-0.0002					352
SARRSDSVI	222	9	-0.0002					353
SARRSDSVIL	222	10	-0.0002					354
SASGHSRT	482	8	-0.0002					355
SASGHSRTT	482	9	-0.0002					356
SASGHSRTTV	482	10	-0.0002					357
SASGTSPGL	675	9	-0.0002					358
SASGTSPGLSA	675	11	0.0001					359
SISSNNSKPV	504	10	-0.0002					360
SITVSASGT	671	9	-0.0002					361
SIVKSITV	667	8	-0.0002					362
SIVKSITVSA	667	10	0.0004					363
SLLIQNII	106	8	0.0008					364
SLLTFWNPPT	23	10	0.0022					365
SLLTFWNPPTT	23	11						366
SLPVSPRL	540	8						367
SLPVSPRLQL	540	10						368
STQELFIPNI	280	10						369
STQELFIPNIT	280	11						370
SVDHSDPV	400	8	0.0001					371
SVDHSDPVI	400	9	-0.0002					372
SVDHSDPVIL	400	10	-0.0002					373
SVSANRSDPV	576	10	-0.0002					374
SVSANRSDPVT	576	11	-0.0001					375
TAKLTIEST	33	9						376
TASYKCET	210	8	0.0001					377
TIESTPFNV	37	9						378
TIESTPFNVA	37	10						379

Table VIII

CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
TITVSAEL	493	8	-0.0002					380
TLDVLYGPDT	586	10	0.0002					381
TLFNVTRNDA	557	10	0.0011					382
TLFNVTRNDT	201	10	0.0003					383
TLFNVTRNDTA	201	11	0.0110					384
TLHVIKSDL	121	9	0.0002					385
TLHVIKSDLV	121	10	0.0017					386
TLLSVTRNDV	379	10	0.0018					387
TLTLFNVT	555	8	0.0001					388
TLTLLSVT	377	8						389
TQDATYLWWV	171	10						390
TQELFIPNI	281	9						391
TQELFIPNIT	281	10						392
TQELFIPNITV	281	11						393
TQELFISNI	459	9						394
TQELFISNIT	459	10						395
TQQATPGPA	86	9						396
TQVLFIAKI	637	9						397
TQVLFIAKIT	637	10						398
TTAKLTIEST	32	10						399
TTVKTITV	489	8	-0.0002					400
TTVKTITVSA	489	10	-0.0002					401
TTVTTITV	311	8	0.0006					402
TTVTTITVYA	311	10	0.0025					403
TVGIMIGV	688	8	0.0004					404
TVGIMIGVL	688	9	0.0014					405
TVGIMIGVLV	688	10	0.0015					406
TVKTITVSA	490	9	-0.0002					407
TVKTITVSAEL	490	11	0.0004					408
TVNNSGSYT	290	9						409
TVSAELPKPSI	495	11	-0.0001					410
TVSASGTSPGL	673	11	-0.0001					411
TVTTITVYA	312	9	0.0047					412
TVYAEPPKPFI	317	11	-0.0001					413
VAEGKEVL	45	8						414
VAEGKEVLL	45	9						415
VAEGKEVLLL	45	10						416
VAEGKEVLLLV	45	11						417
VAFTCEPEA	519	9	0.0011					418
VAFTCEPET	163	9						419
VALTCEPEI	341	9	0.0009					420
VIGTQQAT	83	8						421
VIKSDLVNEEA	124	11						422
VILNVLYGPDA	229	11	0.0001					423
VLFIAKIT	639	8	0.0005					424
VLLLVIINL	51	8						425
VLVGVALI	695	8	0.0073					426
VLYGPDAPT	233	9	0.0030					427
VLYGPDAPTI	233	10	0.0110	0.0130	1.0000	0.0033	0.0016	428
VLYGPDDPT	411	9	0.0005					429

Table VIII

CEA A02 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0201	A*0202	A*0203	A*0206	A*6802	SEQ ID NO.
VLYGPDDPTI	411	10	0.0200	0.0130	0.0720	0.0007	0.0003	430
VLYGPDTPI	589	9	0.0160					431
VLYGPDTPII	589	10	0.0057					432
VTLDVLYGPDT	585	11	-0.0001					433
VTRNDARA	561	8	-0.0002					434
VTRNDARAYV	561	10	0.0002					435
VTTITVYA	313	8	0.0009					436
WLIDGNIQQIIT	449	11	0.0005					437
WQRLLLTA	15	8						438
WQRLLLTASL	15	10						439
WQRLLLTASLL	15	11						440
WVNGQSLPV	535	9	0.0020					441
WVNNQSLPV	357	9	0.0012					442
YACFVSNL	653	8	0.0002					443
YACFVSNLA	653	9	0.0002					444
YACFVSNLAT	653	10	0.0046					445
YAEPPKPFI	319	9	-0.0002					446
YAEPPKPFIT	319	10	-0.0002					447
YLSGANLNL	605	9	0.3600					448
YLWWVNGQSL	532	10	0.1400					449
YLWWVNNQSL	354	10	0.4200					450
YTCQAHNSDT	297	10	-0.0002					451
YTCQANNSA	475	9	-0.0002					452
YTLIIVIKSDL	120	10	0.0023					453
YTLIIVIKSDLV	120	11	0.0083					454
YTYYRPGV	424	8	0.0003					455
YTYYRPGVNL	424	10	0.0018					456
YVCGIQNSV	569	9	0.0260	0.0097	0.0210	0.0300	0.0200	457
YVCGIQNSVSA	569	11	0.0018					458
YVIGTQQA	82	8						459
YVIGTQQAT	82	9						460

Table VIX
CEA A03 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	A*1101	A*3101	A*3301	A*6801	SEQ ID NO.
ASGHSRTTVK	483	10	0.0008	0.0140	0.0002	0.0005	0.0002	461
ASNPSPQYSWR	618	11	0.0016	0.0056				462
ATGRNNSIVK	661	10	0.0017	0.0045				463
ATPGPAYSGR	89	10	0.0004	0.0190	0.0490	0.0180	0.0075	464
DTGFYTLHVIK	116	11	-0.0009	0.0031				465
ELFISNITEK	461	10	0.0028	0.0030				466
ESPSAPPHR	2	9	-0.0002	-0.0001				467
ESTPFNVAEGK	39	11						468
ETQNPVSAR	216	9	0.0011	0.0012				469
ETQNPVSARR	216	10	-0.0002	0.0002				470
FISNITEK	463	8	0.0038	0.0019				471
FVSNLATGR	656	9	0.0019	0.0490	0.0540	0.2800	0.9800	472
GIQNSVSANR	572	10	0.0018	0.0052				473
HLFGYSWYK	61	9	4.9000	2.5000	0.8800	1.6000	2.3000	474
HTQVLFIAK	636	9	0.0093	0.1700	0.1700	0.2200	0.0500	475
ISPLNTSYR	242	9	0.0004	0.0008				476
ISPSYTYYR	420	9	0.0082	0.0420	0.8500	0.0560	0.7100	477
ITVSAELPK	494	9	0.0080	0.1900	0.0002	0.0005	0.0510	478
ITVYAEPPK	316	9	0.0006	0.0170	0.0002	0.0005	0.0610	479
KTITVSAELPK	492	11	0.3600	0.1600	-0.0006	-0.0013	0.0130	480
LATGRNNSIVK	660	11	0.0008	-0.0002				481
LTFWNPPTTAK	25	11						482
LTLFNVTR	556	8	-0.0007	0.0006				483
LTLLSVTR	378	8						484
LVNEEATGQFR	129	11	-0.0009	0.0013				485
NSASGHSR	481	8	0.0040	-0.0004				486
NSDTGLNR	303	8	-0.0004	-0.0004				487
NSKPVEDK	509	8	-0.0007	-0.0001				488
NVTRNDAR	560	8	-0.0004	-0.0004				489
NVTRNDTASYK	204	11	-0.0002	-0.0002				490
PSISSNNSK	503	9	-0.0008	-0.0001				491
PSPQYSWR	621	8	0.0070	0.0009				492
PTISPLNTSYR	240	11	0.0025	0.0041				493
PTISPSYTYYR	418	11	-0.0002	0.1300	0.4100	0.0370	0.1400	494
QAHNSDTGLNR	300	11	-0.0009	-0.0002				495
QANNSASGHSR	478	11	-0.0009	-0.0002				496
QATPGPAYSGR	88	11						497
QSLPVSPR	539	8						498
RLQLSNDNR	368	9	-0.0010	0.0002				499
RLQLSNGNR	546	9	0.0270	0.0013				500
RTLTLFNVTR	554	10	0.1600	1.1000				501
RTLTLLSVTR	376	10	0.0210	0.1100	2.9000	0.0280	0.0500	502
RVYPELPK	139	8	0.0130	0.0440	0.0010	0.0012	0.0004	503
SASGHSRTTVK	482	11	0.0013	0.0006				504
SISSNNSK	504	8	-0.0007	0.0006				505
SSNNSKPVEDK	506	11	-0.0003	0.0004				506
STPFNVAEGK	40	10						507
TISPLNTSYR	241	10	0.0069	0.0380	0.0870	0.0510	1.8000	508
TISPSYTYYR	419	10	0.0032	0.2800	0.2500	0.1700	2.6000	509
TITVSAELPK	493	10	0.0023	0.0490	0.0002	0.0005	0.0250	510

Table VIX
CEA A03 Supermotif with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	A*1101	A*3101	A*3301	A*6801	SEQ ID NO.
TITVYAEPPK	315	10	-0.0005	0.0035				511
TLFNVTRNDAR	557	11	0.0075	0.0003				512
TLTLFNVTR	555	9	0.0021	0.0006				513
TLTLLSVTR	377	9						514
TTITVYAEPPK	314	11	0.0200	0.0280	0.0008	-0.0013	0.3900	515
TVSAELPK	495	8	0.0037	0.0320	-0.0004	0.0012	0.0053	516
TVYAEPPK	317	8	0.0160	0.0220	-0.0004	0.0014	0.0140	517
VSNLATGR	657	8	-0.0009	0.0021				518
VTRNDTASYK	205	10	-0.0009	0.0014				519
YSWYKGER	65	8						520

Table X
CEA A24 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*2401	SEQ ID NO.
ALTCEPEI	342	8		521
ATGQFRVY	134	8		522
ATGQFRVYPEL	134	11		523
ATGRNNSI	661	8		524
ATVGIMIGVL	687	10		525
AVALTCEPEI	340	10		526
AYSGREII	94	8	0.0003	527
AYSGREIIY	94	9		528
CIPWQRLL	12	8		529
CIPWQRLLL	12	9		530
DLVNEEATGQF	128	11		531
DTGFYTLHVI	116	10		532
DVGPYECGI	387	9		533
DVLYGPDTPI	588	10		534
DVLYGPDTPII	588	11		535
EIIYPNASL	99	9		536
EIIYPNASLL	99	10		537
EIIYPNASLLI	99	11		538
EIQNTTYL	348	8		539
EIQNTTYLW	348	9		540
EIQNTTYLWW	348	10		541
ELSVDHSDPVI	398	11		542
ETQDATYL	170	8		543
ETQDATYLW	170	9		544
ETQDATYLWW	170	10		545
EVLLLVHNL	50	9		546
FWNPPTTAKL	27	10	0.0300	547
FYTLHVIKSDL	119	11	0.0250	548
GFYTLHVI	118	8	0.0010	549
GIPQQHTQVL	631	10		550
GIPQQHTQVLF	631	11		551
GLNRTTVTTI	307	10		552
GLSAGATVGI	682	10		553
GLSAGATVGIM	682	11		554
GTFQQSTQEL	275	10		555
GTFQQSTQELF	275	11		556
GTQQATPGPAY	85	11		557
GTYACFVSNL	651	10		558
GVLVGVAL	694	8		559
GVLVGVALI	694	9		560
HLFGYSWY	61	8		561
HTQELFISNI	458	10		562
HTQVLFIAKI	636	10		563
IIQNDTGF	112	8		564
IIQNDTGFY	112	9		565
IIQNDTGFYTL	112	11		566
IISPPDSSY	597	9		567
IISPPDSSYL	597	10		568
IIYPNASL	100	8		569
IIYPNASLL	100	9		570

Table X

CEA A24 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*2401	SEQ ID NO.
IIYPNASLLI	100	10		571
IMIGVLVGVAL	691	11		572
ITEKNSGL	467	8		573
ITEKNSGLY	467	9		574
ITPNNNGTY	645	9		575
ITVNNSGSY	289	9		576
ITVYAEPPKPF	316	11		577
IYPNASLL	101	8	0.0680	578
IYPNASLLI	101	9	6.9000	579
KITPNNNGTY	644	10		580
KLTIESTPF	35	9		581
KTITVSAEL	492	9		582
LLLTASLL	18	8		583
LLLTASLLTF	18	10		584
LLLTASLLTFW	18	11		585
LLLVHNLPQHL	52	11		586
LLTASLLTF	19	9		587
LLTASLLTFW	19	10		588
LLVHNLPQHL	53	10		589
LLVHNLPQHLF	53	11		590
LTASLLTF	20	8		591
LTASLLTFW	20	9		592
LTIESTPF	36	8		593
LVHNLPQHL	54	9		594
LVHNLPQHLF	54	10		595
LVNEEATGQF	129	10		596
LWWVNGQSL	533	9	0.0082	597
LWWVNNQSL	355	9	0.0220	598
LYGPDAPTI	234	9	0.2100	599
LYGPDDPTI	412	9	0.0340	600
LYGPDTPI	590	8	0.0011	601
LYGPDTPII	590	9	0.2600	602
MIGVLVGVAL	692	10		603
MIGVLVGVALI	692	11		604
NIIQNDTGF	111	9		605
NIIQNDTGFY	111	10		606
NIQQHTQEL	454	9		607
NIQQHTQELF	454	10		608
NIQQHTQELFI	454	11		609
NITEKNSGL	466	9		610
NITEKNSGLY	466	10		611
NITVNNSGSY	288	10		612
NLATGRNNSI	659	10		613
NLPQHLFGY	57	9		614
NLPQHLFGYSW	57	11		615
NTSYRSGENL	246	10		616
NVAEGKEVL	44	9		617
NVAEGKEVLL	44	10		618
NVAEGKEVLLL	44	11		619
NVLYGPDAPTI	232	11		620

Table X
CEA A24 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*2401	SEQ ID NO.
NVLYGPDDPTI	410	11		621
NVTRNDARAY	560	10		622
NVTRNDTASY	204	10		623
PFNVAEGKEVL	42	11	-0.0005	624
PIISPPDSSY	596	10		625
PIISPPDSSYL	596	11		626
PTISPLNTSY	240	10		627
PTISPSYTY	418	9		628
PTISPSYTYY	418	10		629
PTTAKLTI	31	8		630
PVEDEDAVAL	334	10		631
PVEDKDAVAF	512	10		632
PVILNVLY	406	8		633
PVSARRSDSVI	220	11		634
PVSPRLQL	542	8		635
PVTLDVLY	584	8		636
PWQRLLLTASL	14	11	0.0370	637
PYECGIQNEL	390	10	0.0002	638
QFRVYPEL	137	8	0.0006	639
QLSNDNRTL	370	9		640
QLSNDNRTLTL	370	11		641
QLSNGNRTL	548	9		642
QLSNGNRTLTL	548	11		643
QVLFIAKI	638	8		644
QYSWFVNGTF	268	10	3.4000	645
QYSWLIDGNI	446	10	0.0150	646
QYSWRINGI	624	9	0.0270	647
RLLLTASL	17	8		648
RLLLTASLL	17	9		649
RLLLTASLLTF	17	11		650
RLQLSNDNRTL	368	11		651
RLQLSNGNRTL	546	11		652
RTTVTTITVY	310	10		653
RVDGNRQI	72	8		654
RVDGNRQII	72	9		655
RVDGNRQIIGY	72	11		656
RVYPELPKPSI	139	11		657
RWCIPWQRL	10	9	0.0130	658
RWCIPWQRLL	10	10	0.0390	659
RWCIPWQRLLL	10	11	0.0790	660
SLLIQNII	106	8		661
SLPVSPRL	540	8		662
SLPVSPRLQL	540	10		663
STQELFIPNI	280	10		664
SVDHSDPVI	400	9		665
SVDHSDPVIL	400	10		666
SVILNVLY	228	8		667
SVTRNDVGPY	382	10		668
SWFVNGTF	270	8	0.0250	669
SWLIDGNI	448	8	0.0005	670

Table X

CEA A24 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*2401	SEQ ID NO.
SYLSGANL	604	8	0.0051	671
SYLSGANLNL	604	10	0.0580	672
SYRSGENL	248	8	-0.0003	673
SYRSGENLNL	248	10	0.0002	674
SYTYYRPGVNL	423	11	0.0550	675
TFQQSTQEL	276	9	0.0012	676
TFQQSTQELF	276	10	0.0160	677
TFQQSTQELFI	276	11	0.0011	678
TFWNPPTTAKL	26	11	0.0026	679
TISPLNTSY	241	9		680
TISPSYTY	419	8		681
TISPSYTYY	419	9		682
TITVSAEL	493	8		683
TLIIVIKSDL	121	9		684
TTVTTITVY	311	9		685
TVGIMIGVL	688	9		686
TVKTITVSAEL	490	11		687
TVNNSGSY	290	8		688
TVSAELPKPSI	495	11		689
TVSASGTSPGL	673	11		690
TVTTITVY	312	8		691
TVYAEPPKPF	317	10		692
TVYAEPPKPFI	317	11		693
TYACFVSNL	652	9	1.2000	694
TYLWWVNGQSL	531	11	0.1300	695
TYLWWVNNQSL	353	11	0.1400	696
TYYRPGVNL	425	9	0.0650	697
TYYRPGVNLSL	425	11	0.0910	698
VLLLVHNL	51	8		699
VLVGVALI	695	8		700
VLYGPDAPTI	233	10		701
VLYGPDDPTI	411	10		702
VLYGPDTPI	589	9		703
VLYGPDTPII	589	10		704
VTRNDARAY	561	9		705
VTRNDTASY	205	9		706
VTRNDVGPY	383	9		707
VYAEPPKPF	318	9	0.2900	708
VYAEPPKPFI	318	10	0.0180	709
VYPELPKPSI	140	10	0.0079	710
WWVNGQSL	534	8	0.0012	711
WWVNNQSL	356	8	0.0009	712
YLSGANLNL	605	9		713
YLWWVNGQSL	532	10		714
YLWWVNNQSL	354	10		715
YTLIIVIKSDL	120	10		716
YTYYRPGVNL	424	10		717
YYRPGVNL	426	8	0.0220	718
YYRPGVNLSL	426	10	0.1400	719

Table XI
CEA B07 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	B*0702	SEQ ID NO.
APPIRWCI	6	8	0.0006	720
APPIRWCIPW	6	10	0.0290	721
APTISPLNTSY	239	11	-0.0002	722
DPTISPSY	417	8	-0.0006	723
DPTISPSYTY	417	10	-0.0002	724
DPTISPSYTYY	417	11	-0.0002	725
DPVILNVL	405	8	-0.0006	726
DPVILNVLY	405	9	-0.0002	727
DPVTLDVL	583	8	-0.0006	728
DPVTLDVLY	583	9	-0.0002	729
EPEAQNTTY	524	9	-0.0002	730
EPEAQNTTYL	524	10	0.0001	731
EPEAQNTTYLW	524	11	-0.0003	732
EPEIQNTTY	346	9	-0.0002	733
EPEIQNTTYL	346	10	0.0001	734
EPEIQNTTYLW	346	11	-0.0003	735
EPETQDATY	168	9	-0.0002	736
EPETQDATYL	168	10	0.0001	737
EPETQDATYLW	168	11	-0.0003	738
GPAYSGREI	92	9	0.2000	739
GPAYSGREIH	92	10	0.0076	740
GPAYSGREIIY	92	11	0.0013	741
GPDAPTISPL	236	10	0.0048	742
GPDDPTISPSY	414	11	-0.0002	743
GPYECGIQNEL	389	11	0.0006	744
IPQQHTQV	632	8	0.0017	745
IPQQHTQVL	632	9	0.1600	746
IPQQHTQVLF	632	10	0.0180	747
IPQQHTQVLFI	632	11	0.0016	748
IPWQRLLL	13	8	0.1100	749
IPWQRLLLTA	13	10	0.0440	750
KPVEDKDA	511	8	-0.0002	751
KPVEDKDAV	511	9	0.0081	752
KPVEDKDAVA	511	10	0.0010	753
KPVEDKDAVAF	511	11	0.0012	754
LPQHLFGY	58	8	-0.0006	755
LPQHLFGYSW	58	10	-0.0002	756
LPQHLFGYSWY	58	11	-0.0002	757
LPVSPRLQL	541	9	0.9100	758
NPPAQYSW	442	8	0.0002	759
NPPAQYSWF	264	9	0.0001	760
NPPAQYSWFV	264	10	0.0013	761
NPPAQYSWL	442	9	0.0051	762
NPPAQYSWLI	442	10	0.0004	763
NPPTTAKL	29	8	0.0005	764
NPPTTAKLTI	29	10	0.0190	765
NPSPQYSW	620	8	-0.0002	766
NPSPQYSWRI	620	10	-0.0002	767
NPVEDEDA	333	8	-0.0002	768
NPVEDEDAV	333	9	0.0001	769

Table XI
CEA B07 Supermotif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	B*0702	SEQ ID NO.
NPVEDEDAVA	333	10	-0.0002	770
NPVEDEDAVAL	333	11	-0.0002	771
NPVSARRSDSV	219	11	-0.0002	772
PPAQYSWF	265	8	0.0011	773
PPAQYSWFV	265	9	0.0001	774
PPAQYSWL	443	8	0.0002	775
PPAQYSWLI	443	9	0.0002	776
PPDSSYLSGA	600	10	-0.0002	777
PPHRWCIPW	7	9	-0.0002	778
PPTTAKLTI	30	9	0.0003	779
RPGVNLSL	428	8	0.0720	780
SPGLSAGA	680	8	0.0008	781
SPGLSAGATV	680	10	0.0027	782
SPPDSSYL	599	8	-0.0006	783
SPPDSSYLSGA	599	11	-0.0003	784
SPQYSWRI	622	8	0.0004	785
SPQYSWRINGI	622	11	0.0043	786
SPSAPPHRW	3	9	0.0013	787
SPSAPPHRWCI	3	11	0.0022	788
SPSYTYYRPGV	421	11	0.0026	789
TPFNVAEGKEV	41	11	0.0007	790
TPGPAYSGREI	90	11	0.0014	791
TPHSPPDSSY	595	11	-0.0002	792
TPNNNGTY	646	8	-0.0006	793
TPNNNGTYA	646	9	0.0011	794
TPNNNGTYACF	646	11	0.0008	795
YPELPKPSI	141	9	0.0120	796
YPNASLLI	102	8	0.0280	797
YPNASLLIQNI	102	11	0.0007	798

Table XII

B27 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
AHNSDTGL	301	8	799
AKITPNNNGTY	643	11	800
AKLTHESTPF	34	10	801
ARAYVCGI	566	8	802
ARRSDSVI	223	8	803
ARRSDSVIL	223	9	804
CHAASNPPAQY	437	11	805
CHSASNPSPQY	615	11	806
DHSDPVIL	402	8	807
DHSDPVILNVL	402	11	808
ERVDGNRQI	71	9	809
ERVDGNRQII	71	10	810
GHSRTTVKTI	485	10	811
GKEVLLLVHNL	48	11	812
GREHYPNASL	97	11	813
GRNNSIVKSI	663	10	814
HRWCIPWQRL	9	10	815
HRWCIPWQRLL	9	11	816
LHVIKSDL	122	8	817
NRQHGYVI	76	9	818
NRSDPVTL	580	8	819
NRSDPVTLDVL	580	11	820
NRTTVTTI	309	8	821
NRTTVTTITVY	309	11	822
PHRWCIPW	8	8	823
PHRWCIPWQRL	8	11	824
QHLFGYSW	60	8	825
QHLFGYSWY	60	9	826
QHTQELFI	457	8	827
QHTQELFISNI	457	11	828
QHTQVLFI	635	8	829
QHTQVLFIAKI	635	11	830
QRLLLTASL	16	9	831
QRLLLTASLL	16	10	832
RRSDSVIL	224	8	833
RRSDSVILNVL	224	11	834
SRTTVKTI	487	8	835
TRNDARAY	562	8	836
TRNDTASY	206	8	837
TRNDVGPY	384	8	838
VHNLPQHL	55	8	839
VHNLPQHLF	55	9	840
VHNLPQHLFGY	55	11	841
VKTITVSAEL	491	10	842
YRPGVNLSL	427	9	843

Table XIII
B58 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
AASNPPAQY	439	9	844
AASNPPAQYSW	439	11	845
ASGIISRTTV	483	9	846
ASGTSPGL	676	8	847
ASLLIQNI	105	8	848
ASLLIQNII	105	9	849
ASNPPAQY	440	8	850
ASNPPAQYSW	440	10	851
ASNPPAQYSWF	262	11	852
ASNPPAQYSWL	440	11	853
ASNPSPQY	618	8	854
ASNPSPQYSW	618	10	855
ASYKCETQNPV	211	11	856
ATGQFRVY	134	8	857
ATGQFRVYPEL	134	11	858
ATGRNNSI	661	8	859
ATGRNNSIV	661	9	860
ATVGIMIGV	687	9	861
ATVGIMIGVL	687	10	862
ATVGIMIGVLV	687	11	863
DAPTISPL	238	8	864
DARAYVCGI	565	9	865
DATYLWWV	173	8	866
DAVALTCEPEI	339	11	867
DSSYLSGANL	602	10	868
DSVILNVL	227	8	869
DSVILNVLY	227	9	870
DTGFYTLHV	116	9	871
DTGFYTLHVI	116	10	872
DTGLNRTTV	305	9	873
EAQNTTYL	526	8	874
EAQNTTYLW	526	9	875
EAQNTTYLWW	526	10	876
EAQNTTYLWWV	526	11	877
EATGQFRV	133	8	878
EATGQFRVY	133	9	879
ESPSAPPIIRW	2	10	880
ETQDATYL	170	8	881
ETQDATYLW	170	9	882
ETQDATYLWW	170	10	883
ETQDATYLWWV	170	11	884
GATVGIMI	686	8	885
GATVGIMIGV	686	10	886
GATVGIMIGVL	686	11	887
GTFQQSTQEL	275	10	888
GTFQQSTQELF	275	11	889
GTQQATPGPAY	85	11	890
GTYACFVSNL	651	10	891
HAASNPPAQY	438	10	892
HSASNPSPQY	616	10	893

Table XIII
B58 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
HSDPVILNV	403	9	894
HSDPVILNVL	403	10	895
HSDPVILNVLY	403	11	896
HSRTTVKTI	486	9	897
HSRTTVKTITV	486	11	898
HTQELFISNI	458	10	899
HTQVLFIAKI	636	10	900
ISNITEKNSGL	464	11	901
ISPLNTSY	242	8	902
ISPPDSSY	598	8	903
ISPPDSSYL	598	9	904
ISPSYTYY	420	8	905
ISSNNSKPV	505	9	906
ITEKNSGL	467	8	907
ITEKNSGLY	467	9	908
ITPNNNGTY	645	9	909
ITSNNSNPV	327	9	910
ITVNNSGSY	289	9	911
ITVYAEPPKPF	316	11	912
KTITVSAEL	492	9	913
LATGRNNSI	660	9	914
LATGRNNSIV	660	10	915
LSAGATVGI	683	9	916
LSAGATVGIM	683	10	917
LSAGATVGIMI	683	11	918
LSGANLNL	606	8	919
LSNDNRTL	371	8	920
LSNDNRTLTL	371	10	921
LSNDNRTLTLL	371	11	922
LSNGNRTL	549	8	923
LSNGNRTLTL	549	10	924
LSNGNRTLTLF	549	11	925
LSVDHSDPV	399	9	926
LSVDHSDPVI	399	10	927
LSVDHSDPVIL	399	11	928
LSVTRNDV	381	8	929
LSVTRNDVGPY	381	11	930
LTASLLTF	20	8	931
LTASLLTFW	20	9	932
LTIESTPF	36	8	933
LTIESTPFNV	36	10	934
LTLLSVTRNDV	378	11	935
NASLLIQNI	104	9	936
NASLLIQNII	104	10	937
NSASGHSRTTV	481	11	938
NSDTGLNRTTV	303	11	939
NSIVKSITV	666	9	940
NSKPVEDKDAV	509	11	941
NSNPVEDEDAV	331	11	942
NSVSANRSDPV	575	11	943

Table XIII
B58 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
NTSYRSGENL	246	10	944
NTTYLWWV	529	8	945
PAQYSWFV	266	8	946
PAQYSWLI	444	8	947
PAYSGREI	93	8	948
PAYSGREII	93	9	949
PAYSGREIIY	93	10	950
PSAPPIRW	4	8	951
PSAPPIRWCI	4	10	952
PSISSNNSKPV	503	11	953
PSPQYSWRI	621	9	954
PSYTYYRPGV	422	10	955
PTISPLNTSY	240	10	956
PTISPSYTY	418	9	957
PTISPSYTYY	418	10	958
PTTAKLTI	31	8	959
QAIINSDTGL	300	9	960
QATPGPAY	88	8	961
QSLPVSPRL	539	9	962
QSLPVSPRLQL	539	11	963
QSTQELFI	279	8	964
QSTQELFIPNI	279	11	965
RAYVCGIQNSV	567	11	966
RSDPVTLDV	581	9	967
RSDPVTLDVL	581	10	968
RSDPVTLDVLY	581	11	969
RSDSVILNV	225	9	970
RSDSVILNVL	225	10	971
RSDSVILNVLY	225	11	972
RSGENLNL	250	8	973
RTLTLFNV	554	8	974
RTLTLLSV	376	8	975
RTTVKTITV	488	9	976
RTTVTTITV	310	9	977
RTTVTTITVY	310	10	978
SAELPKPSI	497	9	979
SAGATVGI	684	8	980
SAGATVGIM	684	9	981
SAGATVGIMI	684	10	982
SANRSDPV	578	8	983
SANRSDPVTL	578	10	984
SAPPIRWCI	5	9	985
SAPPIRWCIPW	5	11	986
SARRSDSV	222	8	987
SARRSDSVI	222	9	988
SARRSDSVIL	222	10	989
SASGHSRTTV	482	10	990
SASGTSPGL	675	9	991
SASNPSPQY	617	9	992
SASNPSPQYSW	617	11	993

Table XIII

B58 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
SSNNSKPV	506	8	994
SSYLSGANL	603	9	995
SSYLSGANLNL	603	11	996
STQELFIPNI	280	10	997
TAKLTIESTPF	33	11	998
TASLLTFW	21	8	999
TSNNSNPV	328	8	1000
TSPGLSAGATV	679	11	1001
TSYRSGENL	247	9	1002
TSYRSGENLNL	247	11	1003
TTVKTITV	489	8	1004
TTVTTITV	311	8	1005
TTVTTITVY	311	9	1006
VAEGKEVL	45	8	1007
VAEGKEVLL	45	9	1008
VAEGKEVLLL	45	10	1009
VAEGKEVLLLV	45	11	1010
VALTCEPEI	341	9	1011
VSAELPKPSI	496	10	1012
VSANRSDPV	577	9	1013
VSANRSDPVTL	577	11	1014
VSARRSDSV	221	9	1015
VSARRSDSVI	221	10	1016
VSARRSDSVIL	221	11	1017
VSASGTSPGL	674	10	1018
VTRNDARAY	561	9	1019
VTRNDARAYV	561	10	1020
VTRNDTASY	205	9	1021
VTRNDVGPY	383	9	1022
YACFVSNL	653	8	1023
YAEPPKPF	319	8	1024
YAEPPKPFI	319	9	1025
YSGREIIY	95	8	1026
YSWFVNGTF	269	9	1027
YSWLIDGNI	447	9	1028
YSWRINGI	625	8	1029
YSWYKGERV	65	9	1030
YTLHVIKSDL	120	10	1031
YTLHVIKSDLV	120	11	1032
YTYYRPGV	424	8	1033
YTYYRPGVNL	424	10	1034

Table XIV
B62 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
ALTCEPEI	342	8	1035
APPHRWCI	6	8	1036
APPHRWCIPW	6	10	1037
APTISPLNTSY	239	11	1038
AQNTTYLW	527	8	1039
AQNTTYLWW	527	9	1040
AQNTTYLWWV	527	10	1041
AQYSWFVNGTF	267	11	1042
AQYSWLIDGNI	445	11	1043
AVALTCEPEI	340	10	1044
DLVNEEATGQF	128	11	1045
DPTISPSY	417	8	1046
DPTISPSYTY	417	10	1047
DPTISPSYTYY	417	11	1048
DPVILNVLY	405	9	1049
DPVTLDVLY	583	9	1050
DVGPYECGI	387	9	1051
DVLYGPDTPI	588	10	1052
DVLYGPDTPII	588	11	1053
EIIYPNASLLI	99	11	1054
EIQNTTYLW	348	9	1055
EIQNTTYLWW	348	10	1056
EIQNTTYLWWV	348	11	1057
ELFIPNITV	283	9	1058
ELSVDIISDPV	398	10	1059
ELSVDIISDPVI	398	11	1060
EPEAQNTTY	524	9	1061
EPEAQNTTYLW	524	11	1062
EPEIQNTTY	346	9	1063
EPEIQNTTYLW	346	11	1064
EPETQDATY	168	9	1065
EPETQDATYLW	168	11	1066
FITSNNSNPV	326	10	1067
FQQSTQELF	277	9	1068
FQQSTQELFI	277	10	1069
GIMIGVLV	690	8	1070
GIMIGVLVGV	690	10	1071
GIPQQIITQV	631	9	1072
GIPQQIITQVLF	631	11	1073
GIQNELSV	394	8	1074
GLNRTTVTTI	307	10	1075
GLSAGATV	682	8	1076
GLSAGATVGI	682	10	1077
GLSAGATVGIM	682	11	1078
GPAYSGREI	92	9	1079
GPAYSGREII	92	10	1080
GPAYSGREIIY	92	11	1081
GPDDPTISPSY	414	11	1082
GVLVGVALI	694	9	1083
HLFGYSWY	61	8	1084

Table XIV
B62 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
IIVIKSDLV	123	8	1085
IIQNDTGF	112	8	1086
IIQNDTGFY	112	9	1087
IISPPDSSY	597	9	1088
IIYPNASLLI	100	10	1089
IMIGVLVGV	691	9	1090
IPQQIITQV	632	8	1091
IPQQIITQVLF	632	10	1092
IPQQIITQVLFI	632	11	1093
IQNDTGFY	113	8	1094
IQNIIQNDTGF	109	11	1095
IQNTTYLW	349	8	1096
IQNTTYLWW	349	9	1097
IQNTTYLWWV	349	10	1098
IQQIITQELF	455	9	1099
IQQIITQELFI	455	10	1100
KITPNNNGTY	644	10	1101
KLTIESTPF	35	9	1102
KLTIESTPFNV	35	11	1103
KPVEDKDAV	511	9	1104
KPVEDKDAVAF	511	11	1105
LLLTASLLTF	18	10	1106
LLLTASLLTFW	18	11	1107
LLSVTRNDV	380	9	1108
LLTASLLTF	19	9	1109
LLTASLLTFW	19	10	1110
LLVIINLPQHLF	53	11	1111
LPQHLFGY	58	8	1112
LPQHLFGYSW	58	10	1113
LPQHLFGYSWY	58	11	1114
LVIINLPQHLF	54	10	1115
LVNEEATGQF	129	10	1116
MIGVLVGV	692	8	1117
MIGVLVGVALI	692	11	1118
NIIQNDTGF	111	9	1119
NIIQNDTGFY	111	10	1120
NIQQIITQELF	454	10	1121
NIQQIITQELFI	454	11	1122
NITEKNSGLY	466	10	1123
NITVNNSGSY	288	10	1124
NLATGRNNSI	659	10	1125
NLATGRNNSIV	659	11	1126
NLPQHLFGY	57	9	1127
NLPQHLFGYSW	57	11	1128
NPPAQYSW	442	8	1129
NPPAQYSWF	264	9	1130
NPPAQYSWFV	264	10	1131
NPPAQYSWLI	442	10	1132
NPPTTAKLTI	29	10	1133
NPSPQYSW	620	8	1134

Table XIV
B62 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
NPSPQYSWRI	620	10	1135
NPVEDEDAV	333	9	1136
NPVSARRSDSV	219	11	1137
NVAEGKEV	44	8	1138
NVLYGPDAPTI	232	11	1139
NVLYGPDDPTI	410	11	1140
NVTRNDARAY	560	10	1141
NVTRNDARAYV	560	11	1142
NVTRNDTASY	204	10	1143
PHSPPDSSY	596	10	1144
PPAQYSWF	265	8	1145
PPAQYSWFV	265	9	1146
PPAQYSWLI	443	9	1147
PPHRWCIPW	7	9	1148
PPTTAKLTI	30	9	1149
PQHLFGYSW	59	9	1150
PQHLFGYSWY	59	10	1151
PQQHTQVLF	633	9	1152
PQQHTQVLFI	633	10	1153
PQYSWRINGI	623	10	1154
PVEDEDAV	334	8	1155
PVEDKDAV	512	8	1156
PVEDKDAVAF	512	10	1157
PVILNVLY	406	8	1158
PVSARRSDSV	220	10	1159
PVSARRSDSVI	220	11	1160
PVTLDVLY	584	8	1161
QQATPGPAY	87	9	1162
QQHTQELF	456	8	1163
QQHTQELFI	456	9	1164
QQHTQVLF	634	8	1165
QQHTQVLFI	634	9	1166
QQSTQELF	278	8	1167
QQSTQELFI	278	9	1168
QVLFIAKI	638	8	1169
RLLLTASLLTF	17	11	1170
RQHGYVI	77	8	1171
RVDGNRQI	72	8	1172
RVDGNRQII	72	9	1173
RVDGNRQHGY	72	11	1174
RVYPELPKPSI	139	11	1175
SISSNNSKPV	504	10	1176
SIVKSITV	667	8	1177
SLLIQNII	106	8	1178
SPGLSAGATV	680	10	1179
SPQYSWRI	622	8	1180
SPQYSWRINGI	622	11	1181
SPSAPPHRW	3	9	1182
SPSAPPHRWCI	3	11	1183
SPSYTYYRPGV	421	11	1184

Table XIV
B62 Supermotif Peptides

Sequence	Position	No. of Amino Acids	SEQ ID NO.
SVDHSDPV	400	8	1185
SVDHSDPVI	400	9	1186
SVILNVLY	228	8	1187
SVSANRSDPV	576	10	1188
SVTRNDVGPY	382	10	1189
TIESTPFNV	37	9	1190
TISPLNTSY	241	9	1191
TISPSYTY	419	8	1192
TISPSYTYY	419	9	1193
TLHVIKSDLV	121	10	1194
TLLSVTRNDV	379	10	1195
TPFNVAEGKEV	41	11	1196
TPGPAYSGREI	90	11	1197
TPHSPPDSSY	595	11	1198
TPNNNGTY	646	8	1199
TPNNNGTYACF	646	11	1200
TQDATYLW	171	8	1201
TQDATYLWW	171	9	1202
TQDATYLWWV	171	10	1203
TQELFIPNI	281	9	1204
TQELFIPNITV	281	11	1205
TQELFISNI	459	9	1206
TQQATPGPAY	86	10	1207
TQVLFIAKI	637	9	1208
TVGIMIGV	688	8	1209
TVGIMIGVLV	688	10	1210
TVNNSGSY	290	8	1211
TVSAELPKPSI	495	11	1212
TVTTITVY	312	8	1213
TVYAEPPKPF	317	10	1214
TVYAEPPKPFI	317	11	1215
VLVGVALI	695	8	1216
VLYGPDAPTI	233	10	1217
VLYGPDDPTI	411	10	1218
VLYGPDTPI	589	9	1219
VLYGPDTPII	589	10	1220
WVNGQSLPV	535	9	1221
WVNNQSLPV	357	9	1222
YPELPKPSI	141	9	1223
YPNASLLI	102	8	1224
YPNASLLIQNI	102	11	1225
YVCGIQNSV	569	9	1226

Table XV
CEA A01 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0101	SEQ ID NO.
ATGQFRVY	134	8	-0.0021	1227
YSGREHY	95	8	0.0150	1228
ISPLNTSY	242	8	-0.0021	1229
ASNPPAQY	262	8	0.0120	1230
ISPSYTYY	420	8	0.0030	1231
ASNPPAQY	440	8	0.0120	1232
ISPPDSSY	598	8	-0.0021	1233
ASNPSPQY	618	8	0.0085	1234
VTRNDTASY	205	9	0.0011	1235
ITVNNSGSY	289	9	0.0100	1236
TTVTTITVY	311	9	0.0011	1237
VTRNDVGPY	383	9	-0.0021	1238
PTISPSYTY	418	9	0.0035	1239
ITEKNSGLY	467	9	0.0390	1240
VTRNDARAY	561	9	0.0011	1241
ITPNNNGTY	645	9	0.0049	1242
DSVILNVLY	227	9	-0.0021	1243
PTISPLNTSY	240	10	0.0250	1244
RTTVTTITVY	310	10	0.0041	1245
PTISPSYTYY	418	10	0.0770	1246
HSASNPSPQY	616	10	0.3400	1247
GTQQATPGPAY	85	11	0.0069	1248
RSDSVILNVLY	225	11	0.5300	1249
LSVTRNDVGPY	381	11	0.0100	1250
HSDPVILNVLY	403	11	0.9700	1251
RSDPVTLDVLY	581	11	3.2000	1252
PEAQNTTY	525	8	-0.0021	1253
TISPSYTY	419	8	0.0038	1254
EPETQDATY	168	9		1255
EPEIQNTTY	346	9		1256
EPEAQNTTY	524	9		1257
QQATPGPAY	87	9	-0.0021	1258
AYSGREHY	94	9	0.0011	1259
TISPLNTSY	241	9	0.0024	1260
AASNPPAQY	261	9	-0.0021	1261
TISPSYTYY	419	9	0.0240	1262
AASNPPAQY	439	9	-0.0021	1263
HSPPDSSY	597	9	0.0021	1264
SASNPSPQY	617	9	0.0031	1265
PDDPTISPSY	415	10	0.0012	1266
EEATGQFRVY	132	10	-0.0017	1267
HAASNPPAQY	260	10	0.0012	1268
HAASNPPAQY	438	10	0.0012	1269
SDSVILNVLY	226	10	0.0041	1270
RVDGNRQIIGY	72	11	0.0850	1271
GPDDPTISPSY	414	11		1272
NEEATGQFRVY	131	11	-0.0017	1273
TCEPETQDATY	166	11	-0.0017	1274
TCEPEIQNTTY	344	11	-0.0017	1275
TCEPEAQNTTY	522	11	0.0017	1276

Table XV

CEA A01 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0101	SEQ ID NO.
GPAYSGREIIY	92	11		1277
CIIAASNPPAQY	259	11	0.0019	1278
CIIAASNPPAQY	437	11	0.0019	1279
CIISASNPSPQY	615	11	0.0026	1280

Table XVI
CEA A03 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	SEQ ID NO.
AASNPPAQY	439	9		1281
ACFVSNLA	654	8		1282
ACFVSNLATGR	654	11		1283
AFTCEPEA	520	8		1284
AFTCEPETQDA	164	11		1285
ASGHSRTTVK	483	10	0.0008	1286
ASGTSPGLSA	676	10		1287
ASNPPAQY	440	8		1288
ASNPPAQYSWF	262	11		1289
ASNPSPQY	618	8		1290
ASNPSPQYSWR	618	11	0.0016	1291
ATGQFRVY	134	8		1292
ATGRNNSIVK	661	10	0.0017	1293
ATPGPAYSGR	89	10	0.0004	1294
AVAFTCEPEA	518	10		1295
CFVSNLATGR	655	10		1296
CGIQNELSVDH	393	11		1297
CGIQNSVSA	571	9		1298
CGIQNSVSANR	571	11		1299
CIPWQRLLLTA	12	11		1300
DAVAFTCEPEA	517	11		1301
DDPTISPSY	416	9		1302
DDPTISPSYTY	416	11		1303
DGNRQHGY	74	9		1304
DLVNEEATGQF	128	11		1305
DSSYLSGA	602	8		1306
DSVILNVLY	227	9		1307
DTGFYTLH	116	8		1308
DTGFYTLHVIK	116	11	-0.0009	1309
EATGQFRVY	133	9		1310
EDKDAVAF	514	8		1311
EGKEVLLLVH	47	10		1312
ELFISNITEK	461	10	0.0028	1313
ESPSAPPH	2	8		1314
ESPSAPPHR	2	9	-0.0002	1315
ESTPFNVA	39	8		1316
ESTPFNVAEGK	39	11		1317
ETQNPVSA	216	8		1318
ETQNPVSAR	216	9	0.0011	1319
ETQNPVSARR	216	10	-0.0002	1320
FGYSWYKGER	63	10		1321
FISNITEK	463	8	0.0038	1322
FTCEPETQDA	165	10		1323
FVSNLATGR	656	9	0.0019	1324
GANLNLSCH	608	9		1325
GANLNLSCHSA	608	11		1326
GFYTLHVIK	118	9		1327
GIMIGVLVGVA	690	11		1328
GIPQQHTQVLF	631	11		1329
GIQNELSVDH	394	10		1330

Table XVI

CEA A03 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	SEQ ID NO.
GIQNSVSA	572	8		1331
GIQNSVSANR	572	10	0.0018	1332
GLYTCQANNSA	473	11		1333
GSYTCQAH	295	8		1334
GTFQQSTQELF	275	11		1335
GTQQATPGPA	85	10		1336
GTQQATPGPAY	85	11		1337
GTSPGLSA	678	8		1338
GTSPGLSAGA	678	10		1339
GTYACFVSNLA	651	11		1340
GVNLSLSCH	430	9		1341
GVNLSLSCHA	430	10		1342
GVNLSLSCHAA	430	11		1343
HAASNPPA	438	8		1344
HAASNPPAQY	438	10		1345
HLFGYSWY	61	8		1346
HLFGYSWYK	61	9	4.9000	1347
HSASNPSPQY	616	10	0.0006	1348
HSDPVILNVLY	403	11		1349
HTQVLFIA	636	8		1350
HTQVLFIAK	636	9	0.0093	1351
IDGNIQQH	451	8		1352
IGTQQATPGPA	84	11		1353
IGVLVGVA	693	8		1354
IGYVIGTQQA	80	10		1355
IIGYVIGTQQA	79	11		1356
IIQNDTGF	112	8		1357
IIQNDTGFY	112	9		1358
IISPPDSSY	597	9		1359
ILNVLYGPDA	230	10		1360
IMIGVLVGVA	691	10	0.0035	1361
ISPLNTSY	242	8		1362
ISPLNTSYR	242	9	0.0004	1363
ISPPDSSY	598	8		1364
ISPSYTYY	420	8		1365
ISPSYTYYR	420	9	0.0082	1366
ITEKNSGLY	467	9		1367
ITPNNNGTY	645	9	0.0008	1368
ITPNNNGTYA	645	10		1369
ITVNNSGSY	289	9	0.0008	1370
ITVSAELPK	494	9	0.0080	1371
ITVYAEPPK	316	9	0.0006	1372
ITVYAEPPKPF	316	11		1373
IVKSITVSA	668	9		1374
KCETQNPVSA	214	10		1375
KCETQNPVSAR	214	11		1376
KGERVDGNR	69	9		1377
KITPNNNGTY	644	10		1378
KITPNNNGTYA	644	11		1379
KLTIESTPF	35	9		1380

Table XVI

CEA A03 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	SEQ ID NO.
KSDLVNEEA	126	9		1381
KTITVSAELPK	492	11	0.3600	1382
LATGRNNSIVK	660	11	0.0008	1383
LFGYSWYK	62	8		1384
LFGYSWYKGER	62	11		1385
LFISNITEK	462	9		1386
LFNVTRNDA	558	9		1387
LFNVTRNDAR	558	10		1388
LFNVTRNDARA	558	11		1389
LFNVTRNDTA	202	10		1390
LIDGNIQQH	450	9		1391
LLLTASLLTF	18	10		1392
LLLVHNLPQH	52	10		1393
LLTASLLTF	19	9	0.0011	1394
LLTFWNPPTTA	24	11		1395
LLVHNLPQH	53	9		1396
LLVHNLPQHLF	53	11		1397
LSCHAASNPPA	435	11		1398
LSGANLNLSCH	606	11		1399
LSLSCHAA	433	8		1400
LSNGNRTLTLF	549	11		1401
LSVTRNDVGPY	381	11		1402
LTASLLTF	20	8		1403
LTFWNPPTTA	25	10		1404
LTFWNPPTTAK	25	11		1405
LTIESTPF	36	8		1406
LTIESTPFNVA	36	11		1407
LTLFNVTR	556	8	-0.0007	1408
LTLFNVTRNDA	556	11		1409
LTLLSVTR	378	8		1410
LVHNLPQH	54	8		1411
LVHNLPQHLF	54	10		1412
LVNEEATGQF	129	10		1413
LVNEEATGQFR	129	11	-0.0009	1414
MIGVLVGVA	692	9		1415
NDTGFYTLH	115	9		1416
NGNRTLTLF	551	9		1417
NGQSLPVSPR	537	10		1418
NHQNDTGF	111	9		1419
NHQNDTGFY	111	10		1420
NIQQHTQELF	454	10		1421
NITEKNSGLY	466	10		1422
NITVNNSGSY	288	10		1423
NLNLSCHA	254	8		1424
NLNLSCHAA	254	9		1425
NLNLSCHSA	610	9		1426
NLPQHLFGY	57	9		1427
NLSLSCHA	432	8		1428
NLSLSCHAA	432	9		1429
NSASGHSR	481	8	0.0040	1430

Table XVI
CEA A03 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	SEQ ID NO.
NSDTGLNR	303	8	-0.0004	1431
NSGLYTCQA	471	9		1432
NSGSYTCQA	293	9		1433
NSGSYTCQAH	293	10		1434
NSIVKSITVSA	666	11		1435
NSKPVEDK	509	8	-0.0007	1436
NSKPVEDKDA	509	10		1437
NSNPVEDEDA	331	10		1438
NVLYGPDA	232	8		1439
NVTRNDAR	560	8	-0.0004	1440
NVTRNDARA	560	9		1441
NVTRNDARAY	560	10		1442
NVTRNDTA	204	8		1443
NVTRNDTASY	204	10		1444
NVTRNDTASYK	204	11	-0.0002	1445
PAYSGREHY	93	10		1446
PDDPTISPSY	415	10		1447
PDSSYLSGA	601	9		1448
PENVAEGK	42	8		1449
PGPAYSGR	91	8		1450
PGVNLSLSCH	429	10		1451
PGVNLSLSCHA	429	11		1452
PHSPPDSSY	596	10		1453
PSISSNNSK	503	9	-0.0008	1454
PSPQYSWR	621	8	0.0070	1455
PTISPLNTSY	240	10	0.0006	1456
PTISPLNTSYR	240	11	0.0025	1457
PTISPSYTY	418	9		1458
PTISPSYTYY	418	10	0.0006	1459
PTISPSYTYYR	418	11	-0.0002	1460
PVEDEDAVA	334	9		1461
PVEDKDAVA	512	9		1462
PVEDKDAVAF	512	10		1463
PVILNVLY	406	8		1464
PVTLDVLY	584	8		1465
QAHNSDTGLNR	300	11	-0.0009	1466
QANNSASGH	478	9		1467
QANNSASGHSR	478	11	-0.0009	1468
QATPGPAY	88	8		1469
QATPGPAYSGR	88	11		1470
QFRVYPELPK	137	10		1471
QSLPVSPR	539	8		1472
RINGIPQQH	628	9	0.1000	1473
RLLLTASLLTF	17	11		1474
RLQLSNDNR	368	9	-0.0010	1475
RLQLSNGNR	546	9	0.0270	1476
RSDPVTLDVLY	581	11		1477
RSDSVILNVLY	225	11		1478
RSGENLNLSCH	250	11		1479
RTLTLFNVTR	554	10	0.1600	1480

Table XVI
CEA A03 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	SEQ ID NO.
RTLTLLSVTR	376	10	0.0210	1481
RTTVKTITVSA	488	11		1482
RTTVTTITVY	310	10	0.0007	1483
RTTVTTITVYA	310	11		1484
RVDGNRQHGY	72	11		1485
RVYPELPK	139	8	0.0130	1486
SASGHSRTTVK	482	11	0.0013	1487
SASGTSPGLSA	675	11		1488
SASNPSPQY	617	9		1489
SCHAASNPPA	436	10		1490
SDLVNEEA	127	8		1491
SDPVILNVLY	404	10		1492
SDPVTLDVLY	582	10		1493
SDSVILNVLY	226	10		1494
SGANLNLSCH	607	10		1495
SGENLNLSCH	251	10		1496
SGENLNLSCHA	251	11		1497
SGHSRTTVK	484	9	0.0006	1498
SGLYTCQA	472	8		1499
SGREHYPNA	96	10		1500
SGSYTCQA	294	8		1501
SGSYTCQAH	294	9	0.0006	1502
SGTSPGLSA	677	9		1503
SGTSPGLSAGA	677	11		1504
SISSNNSK	504	8	-0.0007	1505
SIVKSITVSA	667	10		1506
SSNNSKPVEDK	506	11	-0.0003	1507
STPFNVAEGK	40	10		1508
SVILNVLY	228	8		1509
SVTRNDVGPY	382	10		1510
TAKLTIESTPF	33	11		1511
TCEPEAQNTTY	522	11		1512
TCEPEIQNTTY	344	11		1513
TCEPETQDA	166	9		1514
TCEPETQDATY	166	11		1515
TCQANNSA	476	8		1516
TCQANNSASGH	476	11		1517
TFQQSTQELF	276	10		1518
TFWNPPTTA	26	9		1519
TFWNPPTTAK	26	10	0.0070	1520
TGFYTLHVIK	117	10	0.0005	1521
TGRNNSIVK	662	9		1522
TIESTPFNVA	37	10		1523
TISPLNTSY	241	9		1524
TISPLNTSYR	241	10	0.0069	1525
TISPSYTY	419	8		1526
TISPSYTYY	419	9		1527
TISPSYTYYR	419	10	0.0032	1528
TITVSAELPK	493	10	0.0023	1529
TITVYAEPPK	315	10	0.0005	1530

Table XVI

CEA A03 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*0301	SEQ ID NO.
TLFNVTRNDA	557	10		1531
TLFNVTRNDAR	557	11	0.0075	1532
TLFNVTRNDTA	201	11		1533
TLTLFNVTR	555	9	0.0021	1534
TLTLLSVTR	377	9		1535
TSPGLSAGA	679	9		1536
TTITVYAEPPK	314	11	0.0200	1537
TTVKTITVSA	489	10		1538
TTVTTITVY	311	9	0.0008	1539
TTVTTITVYA	311	10		1540
TVKTITVSA	490	9		1541
TVNNSGSY	290	8		1542
TVSAELPK	495	8	0.0037	1543
TVTTITVY	312	8		1544
TVTTITVYA	312	9		1545
TVYAEPPK	317	8	0.0160	1546
TVYAEPPKPF	317	10	0.0005	1547
VAFTCEPEA	519	9		1548
VCGIQNSVSA	570	10		1549
VDGNRQHGY	73	10		1550
VIKSDLVNEEA	124	11		1551
VILNVLYGPDA	229	11		1552
VLLLVHNLPQH	51	11		1553
VSNLATGR	657	8	-0.0009	1554
VTRNDARA	561	8		1555
VTRNDARAY	561	9	0.0014	1556
VTRNDTASY	205	9	0.0024	1557
VTRNDTASYK	205	10	-0.0009	1558
VTRNDVGPY	383	9		1559
VTTITVYA	313	8		1560
WLIDGNIQQH	449	10		1561
YACFVSNLA	653	9		1562
YAEPPKPF	319	8		1563
YSGREHY	95	8		1564
YSGREHYPNA	95	11		1565
YSWFVNGTF	269	9	0.0011	1566
YSWYKGER	65	8		1567
YTCQANNSA	475	9		1568
YVCGIQNSVSA	569	11		1569
YVIGTQQA	82	8		1570

Table XVII
CEA A11 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*1101	SEQ ID NO.
AASNPPAQY	439	9		1571
ACFVSNLATGR	654	11		1572
ANLNLSCH	609	8		1573
ANNSASGH	479	8		1574
ANNSASGHSR	479	10		1575
ASGHSRTTVK	483	10	0.0140	1576
ASNPPAQY	440	8		1577
ASNPSPQY	618	8		1578
ASNPSPQYSWR	618	11	0.0056	1579
ATGQFRVY	134	8		1580
ATGRNNSIVK	661	10	0.0045	1581
ATPGPAYSGR	89	10	0.0190	1582
CFVSNLATGR	655	10		1583
CGIQNELSVDH	393	11		1584
CGIQNSVSANR	571	11		1585
DDPTISPSY	416	9		1586
DDPTISPSYTY	416	11		1587
DGNRQIIGY	74	9		1588
DSVILNVLY	227	9		1589
DTGFYTLH	116	8		1590
DTGFYTLHVIK	116	11	0.0031	1591
EATGQFRVY	133	9		1592
EGKEVLLLVH	47	10		1593
ELFISNITEK	461	10	0.0030	1594
ENLNLSCH	253	8		1595
ESPSAPPH	2	8		1596
ESPSAPPHR	2	9	-0.0001	1597
ESTPPNVAEGK	39	11		1598
ETQNPVSAR	216	9	0.0012	1599
ETQNPVSARR	216	10	0.0002	1600
FGYSWYKGER	63	10		1601
FISNITEK	463	8	0.0019	1602
FNVTRNDAR	559	9		1603
FNVTRNDARAY	559	11		1604
FNVTRNDTASY	203	11		1605
FVSNLATGR	656	9	0.0490	1606
GANLNLSCH	608	9		1607
GFYTLHVIK	118	9		1608
GIQNELSVDH	394	10		1609
GIQNSVSANR	572	10	0.0052	1610
GNRQIIGY	75	8		1611
GSYTCQAH	295	8		1612
GTQQATPGPAY	85	11		1613
GVNLSLSCH	430	9		1614
HAASNPPAQY	438	10		1615
HLFGYSWY	61	8		1616
HLFGYSWYK	61	9	2.5000	1617
HNLPQHLFGY	56	10		1618
HNSDTGLNR	302	9		1619
HSASNPSPQY	616	10	0.0001	1620

TableXVII

CEA A11 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*1101	SEQ ID NO.
HSDPVILNVLY	403	11		1621
HTQVLFIAK	636	9	0.1700	1622
IDGNIQQH	451	8		1623
IIQNDTGFY	112	9		1624
IISPPDSSY	597	9		1625
INGIPQQH	629	8		1626
ISPLNTSY	242	8		1627
ISPLNTSYR	242	9	0.0008	1628
ISPPDSSY	598	8		1629
ISPSYTYY	420	8		1630
ISPSYTYYR	420	9	0.0420	1631
ITEKNSGLY	467	9		1632
ITPNNNGTY	615	9	0.0001	1633
ITVNNSGSY	289	9	0.0002	1634
ITVSAELPK	494	9	0.1900	1635
ITVYAEPPK	316	9	0.0170	1636
KCETQNPVSAR	214	11		1637
KGERVIDGNR	69	9		1638
KITPNNNGTY	644	10		1639
KTITVSAELPK	492	11	0.1600	1640
LATGRNNSIVK	660	11	-0.0002	1641
LFGYSWYK	62	8		1642
LFGYSWYKGER	62	11		1643
LFISNITEK	462	9		1644
LFNVTRNDAR	558	10		1645
LIDGNIQQH	450	9		1646
LLLVIINLPQH	52	10		1647
LLVIINLPQH	53	9		1648
LSGANLNLSCH	606	11		1649
LSVTRNDVGPY	381	11		1650
LTFWNPPTTAK	25	11		1651
LTLFNVTR	556	8	0.0006	1652
LTLLSVTR	378	8		1653
LVIINLPQH	54	8		1654
LVNEEATGQFR	129	11	0.0013	1655
NDTGFYTLH	115	9		1656
NGQSLPVSPR	537	10		1657
NHQNDTGFY	111	10		1658
NITEKNSGLY	466	10		1659
NITVNNSGSY	288	10		1660
NLPQHLFGY	57	9		1661
NNQSLPVSPR	359	10		1662
NNSASGHSR	480	9		1663
NNSGSYTCQAH	292	11		1664
NNSKPVEDK	508	9		1665
NSASGHSR	481	8	-0.0004	1666
NSDTGLNR	303	8	-0.0004	1667
NSGSYTCQAH	293	10		1668
NSKPVEDK	509	8	-0.0001	1669
NVTRNDAR	560	8	-0.0004	1670

Table XVII

CEA A11 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*1101	SEQ ID NO.
NVTRNDARAY	560	10		1671
NVTRNDTASY	204	10		1672
NVTRNDTASYK	204	11	-0.0002	1673
PAYSGREHY	93	10		1674
PDDPTISPSY	415	10		1675
PENVAEGK	42	8		1676
PGPAYSGR	91	8		1677
PGVNLSLSCH	429	10		1678
PHSPPDSSY	596	10		1679
PNITVNNSGSY	287	11		1680
PSISSNNSK	503	9	-0.0001	1681
PSPQYSWR	621	8	0.0009	1682
PTISPLNTSY	240	10	0.0002	1683
PTISPLNTSYR	240	11	0.0041	1684
PTISPSYTY	418	9		1685
PTISPSYTYY	418	10	0.0018	1686
PTISPSYTYYR	418	11	0.1300	1687
PVILNVLY	406	8		1688
PVTLDVLY	584	8		1689
QAHNSDTGLNR	300	11	-0.0002	1690
QANNSASGH	478	9		1691
QANNSASGHSR	478	11	-0.0002	1692
QATPGPAY	88	8		1693
QATPGPAYSGR	88	11		1694
QERVYPELPK	137	10		1695
QNDTGFYTLH	114	10		1696
QNELSVDH	396	8		1697
QNHQNDTGFY	110	11		1698
QNPVSARR	218	8		1699
QNSVSANR	574	8		1700
QSLPVSPR	539	8		1701
RINGIPQQH	628	9	0.0094	1702
RLQLSNDNR	368	9	0.0002	1703
RLQLSNGNR	546	9	0.0013	1704
RNDTASYK	207	8		1705
RSDPVTLDVLY	581	11		1706
RSDSVILNVLY	225	11		1707
RSGENLNLSCH	250	11		1708
RTLTLFNVTR	554	10	1.1000	1709
RTLTLLSVTR	376	10	0.1100	1710
RTTVTTITVY	310	10	0.0013	1711
RVDGNRQIIGY	72	11		1712
RVYPELPK	139	8	0.0440	1713
SASGHSRTTVK	482	11	0.0006	1714
SASNPSPQY	617	9		1715
SDPVILNVLY	404	10		1716
SDPVTLDVLY	582	10		1717
SDSVILNVLY	226	10		1718
SGANLNLSCH	607	10		1719
SGENLNLSCH	251	10		1720

Table XVII

CEA A11 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*1101	SEQ ID NO.
SGHSRTTVK	484	9	0.0011	1721
SGSYTCQAH	294	9	0.0001	1722
SISSNNSK	504	8	0.0006	1723
SNITEKNSGLY	465	11		1724
SNNSKPVEDK	507	10		1725
SNPSPQYSWR	619	10		1726
SSNNSKPVEDK	506	11	0.0004	1727
STPFNVAEGK	40	10		1728
SVILNVLY	228	8		1729
SVTRNDVGPY	382	10		1730
TCEPEAQNTTY	522	11		1731
TCEPEIQNTTY	344	11		1732
TCEPETQDATY	166	11		1733
TCQANNSASGH	476	11		1734
TFWNPPTTAK	26	10	0.0110	1735
TGFYTLHVIK	117	10	0.0085	1736
TGRNNSIVK	662	9		1737
TISPLNTSY	241	9		1738
TISPLNTSYR	241	10	0.0380	1739
TISPSYTY	419	8		1740
TISPSYTYY	419	9		1741
TISPSYTYYR	419	10	0.2800	1742
TITVSAELPK	493	10	0.0490	1743
TITVYAEPPK	315	10	0.0035	1744
TLFNVTRNDAR	557	11	0.0003	1745
TLTLFNVTR	555	9	0.0006	1746
TLTLLSVTR	377	9		1747
TTITVYAEPPK	314	11	0.0280	1748
TTVTTITVY	311	9	0.0003	1749
TVNNSGSY	290	8		1750
TVSAELPK	495	8	0.0320	1751
TVTTITVY	312	8		1752
TVYAEPPK	317	8	0.0220	1753
VDGNRQIIGY	73	10		1754
VLLLVHNLPQH	51	11		1755
VNEEATGQFR	130	10		1756
VNGQSLPVSPR	536	11		1757
VNLSLSCH	431	8		1758
VNNQSLPVSPR	358	11		1759
VSNLATGR	657	8	0.0021	1760
VTRNDARAY	561	9	0.0002	1761
VTRNDTASY	205	9	0.0002	1762
VTRNDTASYK	205	10	0.0014	1763
VTRNDVGPY	383	9		1764
WLIDGNIQQH	449	10		1765
WNPPTTAK	28	8		1766
YSGREHY	95	8		1767
YSWYKGER	65	8		1768

TableXVIII

CEA A24 Motif Peptides with Binding Data

Sequence	Position	No. of Amino Acids	A*2401	SEQ ID NO.
AYSGREH	94	8	0.0003	1769
FWNPPTTAKL	27	10	0.0300	1770
FYTLHVIKSDL	119	11	0.0250	1771
GFYTLHVI	118	8	0.0010	1772
IMIGVLVGVAL	691	11		1773
IYPNASLL	101	8	0.0680	1774
IYPNASLLI	101	9	6.9000	1775
LWWVNGQSL	533	9	0.0082	1776
LWWVNNQSL	355	9	0.0220	1777
LYGPDAPTI	234	9	0.2100	1778
LYGPDDPTI	412	9	0.0340	1779
LYGPDTPI	590	8	0.0011	1780
LYGPDTPII	590	9	0.2600	1781
PFNVAEGKEVL	42	11	-0.0005	1782
PWQRLLLTASL	14	11	0.0370	1783
PYECGIQNEL	390	10	0.0002	1784
QFRVYPEL	137	8	0.0006	1785
QYSWFVNGTF	268	10	3.4000	1786
QYSWLIDGNI	446	10	0.0150	1787
QYSWRINGI	624	9	0.0270	1788
RWCIPWQRL	10	9	0.0130	1789
RWCIPWQRLL	10	10	0.0390	1790
RWCIPWQRLLL	10	11	0.0790	1791
SWFVNGTF	270	8	0.0250	1792
SWLIDGNI	448	8	0.0005	1793
SYLSGANL	604	8	0.0051	1794
SYLSGANLNL	604	10	0.0580	1795
SYRSGENL	248	8	-0.0003	1796
SYRSGENLNL	248	10	0.0002	1797
SYTYYRPGVNL	423	11	0.0550	1798
TFQQSTQEL	276	9	0.0012	1799
TFQQSTQELF	276	10	0.0160	1800
TFQQSTQELFI	276	11	0.0011	1801
TFWNPPTTAKL	26	11	0.0026	1802
TYACFVSNL	652	9	1.2000	1803
TYLWWVNGQSL	531	11	0.1300	1804
TYLWWVNNQSL	353	11	0.1400	1805
TYYRPGVNL	425	9	0.0650	1806
TYYRPGVNLSL	425	11	0.0910	1807
VYAEPPKPF	318	9	0.2900	1808
VYAEPPKPFI	318	10	0.0180	1809
VYPELPKPSI	140	10	0.0079	1810
WWVNGQSL	534	8	0.0012	1811
WWVNNQSL	356	8	0.0009	1812
YYRPGVNL	426	8	0.0220	1813
YYRPGVNLSL	426	10	0.1400	1814

Table XIX

CEA DR Super Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	Position	DR1	DR2wβ1	DR2w2β2	DR3	DR4w4	DR4w15	DR5w11	DR5w12	SEQ ID NO.
IPWQRLLLT	RWCIPWQRLLLTASL	10	0.6100	0.0110	-0.0007	0.0150	0.0830		-0.0005		1815
WQRLLLTAS	CIPWQRLLLTASLLT	12									1816
LLLTASLLT	WQRLLLTASLLTFWN	15									1817
LLTASLLTF	QRLLLTASLLTFWNP	16	-0.0004				-0.0022				1818
LTASLLTFW	RLLLTASLLTFWNPP	17									1819
LTFWNPPTT	ASLLTFWNPPTTAKL	22									1820
FWNPPTTAK	LLTFWNPPTTAKLTI	24									1821
WNPPTTAKL	LTFWNPPTTAKLTIE	25									1822
LTIESTPFN	TAKLTIESTPFNVAE	33									1823
LLVHNLPQH	EVLLLVHNLPQHLFG	50	2.5000	0.2300	0.0013	0.8900	0.8600		0.0340		1824
LVHNLPQHL	VLLLVHNLPQHLFGY	51									1825
YKGERVDGN	YSWYKGERVDGNRQI	65									1826
IIGYVIGTQ	NRQIIGYVIGTQQAT	76									1827
IGTQQATPG	GYVIGTQQATPGPAY	81									1828
YSGREIIYP	GPAYSGREIIYPNAS	92									1829
IIYPNASLL	GREIIYPNASLLIQN	97	0.6200	0.3800	0.0024	0.2700	0.0930		0.0029		1830
IYPNASLLI	REIIYPNASLLIQNI	98									1831
YPNASLLIQ	EIIYPNASLLIQNII	99	0.3500	0.1600	-0.0007	0.1400	0.0390		-0.0005		1832
LLIQNIIQN	NASLLIQNIIQNDTG	104	0.0011				-0.0022				1833
LIQNIIQND	ASLLIQNIIQNDTGF	105									1834
IIQNDTGFY	IQNIIQNDTGFYTLII	109									1835
FYTLHVIKS	DTGFYTLHVIKSDLV	116	0.0720	0.0180	0.0250	0.0013	0.0260		0.0080		1836
YTLHVIKSD	TGFYTLHVIKSDLVN	117									1837
LHVIKSDLV	FYTLHVIKSDLVNEE	119									1838
VIKSDLVNE	TLHVIKSDLVNEEAT	121									1839
IKSDLVNEE	LHVIKSDLVNEEATG	122				0.1300					1840
LVNEEATGQ	KSDLVNEEATGQFRV	126				0.0058					1841
VNEEATGQF	SDLVNEEATGQFRVY	127				-0.0027					1842
VYPELPKPS	QFRVYPELPKPSISS	137				-0.0027					1843
LPKPSISSN	YPELPKPSISSNNSK	141	0.0009				-0.0022				1844
ISSNNSKPV	KPSISSNNSKPVEDK	146	0.0021				-0.0022				1845
VEDKDAVAF	SKPVEDKDAVAFTCE	154									1846
WVNNQSLPV	YLWWVNNQSLPVSPR	176	8.4000	0.0830	0.0095	0.1300	5.6000		0.7000		1847
VNNQSLPVS	LWWVNNQSLPVSPRL	177	0.0230				0.0290				1848
LTLFNVTRN	NRTLTLFNVTRNDTA	197									1849
VTRNDTASY	LFNVTRNDTASYKCE	202									1850
VSARRSDSV	QNPVSARRSDSVILN	218									1851
VILNVLYGP	SDSVILNVLYGPDAP	226									1852
LYGPDAPTI	LNVLYGPDAPTISPL	231									1853
YGPDAPTIS	NVLYGPDAPTISPLN	232				-0.0027					1854
ISPLNTSYR	APTISPLNTSYRSGE	239									1855
LSCHAASNP	NLNLSCHAASNPPAQ	254									1856
WFVNGTFQQ	QYSWFVNGTFQQSTQ	268	0.0260	-0.0007	0.0033	0.0280	0.5600		0.0540		1857
LFIPNITVN	TQELFIPNITVNNSG	281									1858
FIPNITVNN	QELFIPNITVNNSGS	282									1859
IPNITVNNS	ELFIPNITVNNSGSY	283									1860
ITVNNSGSY	IPNITVNNSGSYTCQ	286									1861
VNNSGSYTC	NITVNNSGSYTCQAII	288									1862
LNRTTVTTI	DTGLNRTTVTTITVY	305	-0.0004				-0.0022				1863
VTTITVYAE	RTTVTTITVYAEPPK	310									1864

Table XIX

CEA DR Super Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	DR6w19	DR7	DR8w2	DR9	DRw53	SEQ ID NO.
IPWQRLLLT	RWCIPWQRLLLTASL	0.0110	0.0700	-0.0004			1815
WQRLLLTAS	CIPWQRLLLTASLLT						1816
LLLTASLLT	WQRLLLTASLLTFWN						1817
LLTASLLTF	QRLLLTASLLTFWNP		-0.0013				1818
LTASLLTFW	RLLLTASLLTFWNPP						1819
LTFWNPPTT	ASLLTFWNPPTTAKL						1820
FWNPPTTAK	LLTFWNPPTTAKLTI						1821
WNPPTTAKL	LTFWNPPTTAKLTIE						1822
LTIESTPFN	TAKLTIESTPFNVAE						1823
LLVHNLPQH	EVLLLVHNLPQHLFG	3.4000	0.4700	0.1200			1824
LVHNLPQHL	VLLLVHNLPQHLFGY						1825
YKGERVDGN	YSWYKGERVDGNRQI						1826
IIGYVIGTQ	NRQIIGYVIGTQQAT						1827
IGTQQATPG	GYVIGTQQATPGPAY						1828
YSGREHYP	GPAYSGREHYPNAS						1829
HYPNASLL	GREHYPNASLLIQN	1.2000	0.5600	0.0083			1830
IYPNASLLI	REHYPNASLLIQNI						1831
YPNASLLIQ	EHYPNASLLIQNII	0.3100	0.1600	0.0029			1832
LLIQNIIQN	NASLLIQNIIQNDTG		-0.0013				1833
LIQNIIQND	ASLLIQNIIQNDTGF						1834
IIQNDTGFY	IQNIIQNDTGFYTLH						1835
FYTLHVIKS	DTGFYTLHVIKSDLV	0.0009	0.1100	0.0620			1836
YTLHVIKSD	TGFYTLHVIKSDLVN						1837
LHVIKSDLV	FYTLHVIKSDLVNEE						1838
VIKSDLVNE	TLHVIKSDLVNEEAT						1839
IKSDLVNEE	LHVIKSDLVNEEATG						1840
LVNEEATGQ	KSDLVNEEATGQFRV						1841
VNEEATGQF	SDLVNEEATGQFRVY						1842
VYPELPKPS	QFRVYPELPKPSISS						1843
LPKPSISSN	YPELPKPSISSNNSK		-0.0013				1844
ISSNNSKPV	KPSISSNNSKPVEDK		0.0033				1845
VEDKDAVAF	SKPVEDKDAVAFTCE						1846
WVNNQSLPV	YLWWVNNQSLPVSPR	1.5000	0.6000	0.0460			1847
VNNQSLPVS	LWWVNNQSLPVSPRL		0.0082				1848
LTLFNVTRN	NRTLTLFNVTRNDTA						1849
VTRNDTASY	LFNVTRNDTASYKCE						1850
VSARRSDSV	QNPVSARRSDSVILN						1851
VILNVLYGP	SDSVILNVLYGPDAP						1852
LYGPDAPTI	LNVLYGPDAPTISPL						1853
YGPDAPTIS	NVLYGPDAPTISPLN						1854
ISPLNTSYR	APTISPLNTSYRSGE						1855
LSCHAASNP	NLNLSCHAASNPPAQ						1856
WFVNGTFQQ	QYSWFVNGTFQQSTQ	0.0006	0.0270	0.0039			1857
LFIPNITVN	TQELFIPNITVNNSG						1858
FIPNITVNN	QELFIPNITVNNSGS						1859
IPNITVNNS	ELFIPNITVNNSGSY						1860
ITVNNSGSY	IPNITVNNSGSYTCQ						1861
VNNSGSYTC	NITVNNSGSYTCQAH						1862
LNRTTVTTI	DTGLNRTTVTTITVY		0.0088				1863
VTTITVYAE	RTTVTTITVYAEPPK						1864

Table XIX

CEA DR Super Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	Position	DR1	DR2wβ1	DR2wβ2	DR3	DR4w4	DR4w15	DR5w11	DR5w12	SEQ ID NO.
VYAEPPKPF	TITVYAEPPKPFITS	315				0.0042					1865
ITSNNSNPV	KPFITSNNSNPVEDE	324	-0.0004				-0.0022				1866
VEDEDAVAL	SNPVEDEDAVALTCE	332				0.0054					1867
LTLLSVTRN	NRTLTLLSVTRNDVG	375	0.0210				-0.0022				1868
VTRNDVGPY	LLSVTRNDVGPYECG	380									1869
VGPYECGIQ	RNDVGPYECGIQNEL	385									1870
IQNELSVDH	ECGIQNELSVDHSDP	392				-0.0027					1871
LSVDHSDPV	QNELSVDHSDPVILN	396				0.0820					1872
VDHSDPVIL	ELSVDHSDPVILNVL	398									1873
VILNVLYGP	SDPVILNVLYGPDDP	404									1874
YGPDDPTIS	NVLYGPDDPTISPSY	410				-0.0027					1875
ISPSYTYYR	DPTISPSYTYYRPGV	417									1876
YTYYRPGVN	SPSYTYYRPGVNLSL	421									1877
YYRPGVNLS	SYTYYRPGVNLSLSC	423									1878
VNLSLSCHA	RPGVNLSLSCHAASN	428									1879
LSCHAASNP	NLSLSCHAASNPPAQ	432									1880
LIDGNIQQH	YSWLIDGNIQQHTQE	447									1881
LFISNITEK	TQELFISNITEKNSG	459									1882
FISNITEKN	QELFISNITEKNSGL	460	0.0005				0.0180				1883
ITEKNSGLY	ISNITEKNSGLYTCQ	464									1884
LYTCQANNS	NSGLYTCQANNSASG	471									1885
VKTITVSAE	RTTVKTITVSAELPK	488	0.0110	0.0250	0.0009	0.0010	0.0064		-0.0005		1886
VSAELPKPS	TITVSAELPKPSISS	493				-0.0027					1887
LPKPSISSN	SAELPKPSISSNNSK	497	-0.0004				-0.0022				1888
WVNGQSLPV	YLWWVNGQSLPVSPR	532									1889
VNGQSLPVS	LWWVNGQSLPVSPRL	533									1890
LTLFNVTRN	NRTLTLFNVTRNDAR	553									1891
VTRNDARAY	LFNVTRNDARAYVCG	558									1892
IQNSVSANR	VCGIQNSVSANRSDP	570									1893
VSANRSDPV	QNSVSANRSDPVTLD	574									1894
VTLDVLYGP	SDPVTLDVLYGPDTP	582				-0.0027					1895
LYGPDTPII	LDVLYGPDTPIISPP	587	-0.0004				-0.0022				1896
YGPDTPIIS	DVLYGPDTPIISPPD	588				0.0037					1897
ISPPDSSYL	TPIISPPDSSYLSGA	595									1898
LSGANLNLS	SSYLSGANLNLSCHS	603									1899
LSCHSASNP	NLNLSCHSASNPSPQ	610									1900
WRINGIPQQ	QYSWRINGIPQQHTQ	624									1901
IPQQHTQVL	INGIPQQHTQVLFIA	629									1902
LFIAKITPN	TQVLFIAKITPNNNG	637	0.0820				0.0037				1903
FIAKITPNN	QVLFIAKITPNNNGT	638	0.1200				0.0240				1904
IAKITPNNN	VLFIAKITPNNNGTY	639									1905
YACFVSNLA	NGTYACFVSNLATGR	650									1906
FVSNLATGR	YACFVSNLATGRNNS	653	0.0240				0.0270				1907
VSNLATGRN	ACFVSNLATGRNNSI	654									1908
IVKSITVSA	NNSIVKSITVSASGT	665	0.0550	0.0029	-0.0007	0.1100	1.8000		0.0016		1909
VKSITVSAS	NSIVKSITVSASGTS	666	0.0640	0.0023	-0.0007	0.0750	1.8000		0.0012		1910
ITVSASGTS	VKSITVSASGTSPGL	669									1911
VSASGTSPG	SITVSASGTSPGLSA	671									1912
LSAGATVGI	SPGLSAGATVGIMIG	680									1913
IMIGVLVGV	TVGIMIGVLVGVALI	688									1914

Table XIX

CEA DR Super Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	DR6w19	DR7	DR8w2	DR9	DRw53	SEQ ID NO.
VYAEPPKPF	TITVYAEPPKPFITS						1865
ITSNNSNPV	KPFITSNNSNPVEDE		-0.0013				1866
VEDEDAVAL	SNPVEDEDAVALTCE						1867
LTLLSVTRN	NRTLTLLSVTRNDVG		0.0021				1868
VTRNDVGPY	LLSVTRNDVGPYECG						1869
VGPYECGIQ	RNDVGPYECGIQNEL						1870
IQNELSVDH	ECGIQNELSVDHSDP						1871
LSVDHSDPV	QNELSVDHSDPVILN						1872
VDHSDPVIL	ELSVDHSDPVILNVL						1873
VILNVLYGP	SDPVILNVLYGPDDP						1874
YGPDDPTIS	NVLYGPDDPTISPSY						1875
ISPSYTYYR	DPTISPSYTYYRPGV						1876
YTYYRPGVN	SPSYTYYRPGVNLSL						1877
YYRPGVNLS	SYTYYRPGVNLSLSC						1878
VNLSLSCHA	RPGVNLSLSCHAASN						1879
LSCHAASNP	NLSLSCHAASNPPAQ						1880
LIDGNIQQH	YSWLIDGNIQQHTQE						1881
LFISNITEK	TQELFISNITEKNSG						1882
FISNITEKN	QELFISNITEKNSGL		-0.0013				1883
ITEKNSGLY	ISNITEKNSGLYTCQ						1884
LYTCQANNS	NSGLYTCQANNSASG						1885
VKTITVSAE	RTTVKTITVSAELPK	0.0050	0.0790	-0.0004			1886
VSAELPKPS	TITVSAELPKPSISS						1887
LPKPSISSN	SAELPKPSISSNNSK		-0.0013				1888
WVNGQSLPV	YLWWVNGQSLPVSPR						1889
VNGQSLPVS	LWWVNGQSLPVSPRL						1890
LTLFNVTRN	NRTLTLFNVTRNDAR						1891
VTRNDARAY	LFNVTRNDARAYVCG						1892
IQNSVSANR	VCGIQNSVSANRSDP						1893
VSANRSDPV	QNSVSANRSDPVTLD						1894
VTLDVLYGP	SDPVTLDVLYGPDTP						1895
LYGPDTPII	LDVLYGPDTPIISPP		-0.0013				1896
YGPDTPIIS	DVLYGPDTPIISPPD						1897
ISPPDSSYL	TPIISPPDSSYLSGA						1898
LSGANLNLS	SSYLSGANLNLSCHS						1899
LSCHSASNP	NLNLSCHSASNPSPQ						1900
WRINGIPQQ	QYSWRINGIPQQHTQ						1901
IPQQHTQVL	INGIPQQHTQVLFIA						1902
LFIAKITPN	TQVLFIAKITPNNNG		0.0038				1903
FIAKITPNN	QVLFIAKITPNNNGT		0.0024				1904
IAKITPNNN	VLFIAKITPNNNGTY						1905
YACFVSNLA	NGTYACFVSNLATGR						1906
FVSNLATGR	YACFVSNLATGRNNS		0.0070				1907
VSNLATGRN	ACFVSNLATGRNNSI						1908
IVKSITVSA	NNSIVKSITVSASGT	0.0690	0.0370	0.0120			1909
VKSITVSAS	NSIVKSITVSASGTS	0.0460	0.0760	0.0170			1910
ITVSASGTS	VKSITVSASGTSPGL						1911
VSASGTSPG	SITVSASGTSPGLSA						1912
LSAGATVGI	SPGLSAGATVGIMIG						1913
IMIGVLVGV	TVGIMIGVLVGVALI						1914

Table XIX

CEA DR Super Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	Position	DR1	DR2wβ1	DR2w2β2	DR3	DR4w4	DR4w15	DR5w11	DR5w12	SEQ ID NO.
LTIESTPFN	TAKLTIESTPFNVAE	33									1915
YKGERVDGN	YSWYKGERVDGNRQI	65									1916
LPVSPRLQL	NQSLPVSPRLQLSNG	182									1917
LNLSCHAAS	GENLNLSCHAASNPP	252									1918
LPVSPRLQL	GQSLPVSPRLQLSNG	538									1919

TableXIX

CEA DR Super Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	DR6w19	DR7	DR8w2	DR9	DRw53	SEQ ID NO.
LTIESTPFN	TAKLTIESTPFNVAE						1915
YKGERVDGN	YSWYKGERVDGNRQI						1916
LPVSPRLQL	NQSLPVSPRLQLSNG						1917
LNLSCHAAS	GENLNLSCHAASNPP						1918
LPVSPRLQL	GQSLPVSPRLQLSNG						1919

Table XXa

CEA DR 3a Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	Position	DR1	DR2w2B1	DR2w2B2	DR3	DR4w4	DR4w15	DR5w11	DR5w12	SEQ ID NO.
IQNDTGFYT	QNIIQNDTGFYTLHV	110	0.0044	0.0105	-0.0007	0.3200	-0.0055		-0.0008		1920
IKSDLVNEE	LHVIKSDLVNEEATG	122				0.1300					1921
LVNEEATGQ	KSDLVNEEATGQFRV	126				0.0058					1922
VNEEATGQF	SDLVNEEATGQFRVY	127				-0.0027					1923
VYPELPKPS	QFRVYPELPKPSISS	137				-0.0027					1924
FTCEPETQD	AVAFTCEPETQDATY	162				-0.0027					1925
YKCETQNPV	TASYKCETQNPVSAR	210				-0.0027					1926
YGPDAPTIS	NVLYGPDAPTISPLN	232				-0.0027					1927
VYAEPPKPF	TITVYAEPPKPFITS	315				0.0042					1928
VEDEDAVAL	SNPVEDEDAVALTCE	332				0.0054					1929
LTCEPEIQN	AVALTCEPEIQNTTY	340				0.0039					1930
IQNELSVDH	ECGIQNELSVDHSDP	392				-0.0027					1931
LSVDHSDPV	QNELSVDHSDPVILN	396				0.0820					1932
YGPDDPTIS	NVLYGPDDPTISPSY	410				-0.0027					1933
VSAELPKPS	TITVSAELPKPSISS	493				-0.0027					1934
FTCEPEAQN	AVAFTCEPEAQNTTY	518				-0.0027					1935
VTLDVLYGP	SDPVTLDVLYGPDTP	582				-0.0027					1936
YGPDTPIIS	DVLYGPDTPIISPPD	588				0.0037					1937

Table XXa

CEA DR 3a Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	DR6w19	DR7	DR8w2	DR9	DRw53	SEQ ID NO.
							1920
IQNDTGFYT	QNIIQNDTGFYTLHV	0.3600	-0.0017	-0.0009			1921
IKSDLVNEE	LHVIKSDLVNEEATG						1922
LVNEEATGQ	KSDLVNEEATGQFRV						1923
VNEEATGQF	SDLVNEEATGQFRVY						1924
VYPELPKPS	QFRVYPELPKPSISS						1925
FTCEPETQD	AVAFTCEPETQDATY						1926
YKCETQNPV	TASYKCETQNPVSAR						1927
YGPDAPTIS	NVLYGPDAPTISPLN						1928
VYAEPPKPF	TITVYAEPPKPFITS						1929
VEDEDAVAL	SNPVEDEDAVALTCE						1930
LTCEPEIQN	AVALTCEPEIQNTTY						1931
IQNELSVDH	ECGIQNELSVDHSDP						1932
LSVDHSDPV	QNELSVDHSDPVILN						1933
YGPDDPTIS	NVLYGPDDPTISPSY						1934
VSAELPKPS	TITVSAELPKPSISS						1935
FTCEPEAQN	AVAFTCEPEAQNTTY						1936
VTLDVLYGP	SDPVTLDVLYGPDTP						1937
YGPDTPIIS	DVLYGPDTPIISPPD						

Table XXb

CEA DR 3b Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	Position	DR1	DR2w201	DR2w202	DR3	DR4w4	DR4w15	DR5w11	DR5w12	SEQ ID NO.
ATGQFRVYP	NEEATGQFRVYPELP	131				-0.0027					1938
LNTSYRSGE	ISPLNTSYRSGENLN	242				-0.0027					1939
YTCQAHNSD	SGSYTCQAHNSDTGL	294				-0.0027					1940
LPVSPRLQL	NQSLPVSPRLQLSND	360				0.0071					1941
LSNDNRTLT	RLQLSNDNRTLTLLS	368	0.0001	-0.0006	-0.0007	0.3200	-0.0055		-0.0008		1942
LSLSCHAAS	GVNLSLSCHAASNPP	430				0.0075					1943
LNLSCHSAS	GANLNLSCHSASNPS	608				-0.0027					1944
ASPETHLDM	RLPASPETHLDMLRH	34				-0.0027					1945
AHNQVRQVP	VLIAHNQVRQVPLQR	84				0.0290					1946
LIDTNRSRA	ALTLIDTNRSRACHP	180				0.0350					1947
IHHNTHLCF	LALIHHNTHLCFVHT	465	0.0140	0.0990	0.0009	0.3100	-0.0055		0.0025		1948
LFRNPHQAL	WDQLFRNPHQALLHT	482	-0.0001	0.0015	-0.0007	0.9000	-0.0055		-0.0008		1949
VDLDDKGCP	HSCVDLDDKGCPAEQ	632				-0.0027					1950
YLEDVRLVH	GMSYLEDVRLVHRDL	832				0.1800					1951
IDSECRPRF	CWMIDSECRPRFREL	958	0.0036	-0.0006	0.0150	0.4500	-0.0055		-0.0008		1952
AAPQPHPPP	QGGAAPQPHPPPAFS	1200				-0.0025					1953
AAISRKMVE	EFQAAISRKMVELVH	104				0.0039					1954
LHHTLKIGG	VKVLHHTLKIGGEPH	284				-0.0025					1955
IGGEPHISY	TLKIGGEPHISYPPL	290				-0.0025					1956
AALSRKVAE	EFQAALSRKVAELVH	104				0.0027					1957
ILGDPKKLL	EDSILGDPKKLLTQH	235	0.0003	-0.0006	-0.0010	0.6700	-0.0055		-0.0008		1958
YKQSQHMTE	MAIYKQSQHMTEVVR	160				-0.0025					1959
VEGNLRVEY	LIRVEGNLRVEYLDD	194				0.0930					1960
FTLQIRGRE	GEYFTLQIRGRERFE	325				0.0290					1961

Table XXb

CEA DR 3b Motif Peptides with Binding Data

Core Sequence	Exemplary Sequence	DR6w19	DR7	DR8w2	DR9	DRw53	SEQ ID NO.
							1938
ATGQFRVYP	NEEATGQFRVYPELP						1939
LNTSYRSGE	ISPLNTSYRSGENLN						1940
YTCQAHNSD	SGSYTCQAHNSDTGL						1941
LPVSPRLQL	NQSLPVSPRLQLSND						1942
LSNDNRTLT	RLQLSNDNRTLTLLS	0.0048	-0.0017	-0.0009			1943
LSLSCHAAS	GVNLSLSCHAASNPP						1944
LNLSCHSAS	GANLNLSCHSASNPS						1945
ASPETHLDM	RLPASPETHLDMLRH						1946
AHNQVRQVP	VLIAHNQVRQVPLQR						1947
LIDTNRSRA	ALTLIDTNRSRACHP						1948
IIHNTHLCF	LALIHHNTHLCFVHT	0.7500	0.0200	0.0330			1949
LFRNPHQAL	WDQLFRNPHQALLHT	0.0410	-0.0017	-0.0009			1950
VDLDDKGCP	HSCVDLDDKGCPAEQ						1951
YLEDVRLVH	GMSYLEDVRLVHRDL						1952
IDSECRPRF	CWMIDSECRPRFREL	(0.0001)	-0.0014	0.0028			1953
AAPQPHPPP	QGGAAPQPHPPPAFS						1954
AAISRKMVE	EFQAAISRKMVELVH						1955
LHHTLKIGG	VKVLHHTLKIGGEPH						1956
IGGEPHISY	TLKIGGEPHISYPPL						1957
AALSRKVAE	EFQAALSRKVAELVH						1958
ILGDPKKLL	EDSILGDPKKLLTQH	0.0130	-0.0014	0.0029			1959
YKQSQHMTE	MAIYKQSQHMTEVVR						1960
VEGNLRVEY	LIRVEGNLRVEYLDD						1961
FTLQIRGRE	GEYFTLQIRGRERFE						

TABLE XXI. Population coverage with combined HLA Supertypes

	PHENOTYPIC FREQUENCY					
HLA-SUPERTYPES	Caucasian	North American Black	Japanese	Chinese	Hispanic	Average
a. Individual Supertypes						
A2	45.8	39.0	42.4	45.9	43.0	43.2
A3	37.5	42.1	45.8	52.7	43.1	44.2
B7	38.6	52.7	48.8	35.5	47.1	44.7
A1	47.1	16.1	21.8	14.7	26.3	25.2
A24	23.9	38.9	58.6	40.1	38.3	40.0
B44	43.0	21.2	42.9	39.1	39.0	37.0
B27	28.4	26.1	13.3	13.9	35.3	23.4
B62	12.6	4.8	36.5	25.4	11.1	18.1
B58	10.0	25.1	1.6	9.0	5.9	10.3
b. Combined Supertypes						
A2, A3, B7	83.0	86.1	87.5	88.4	86.3	86.2
A2, A3, B7, A24, B44, A1	99.5	98.1	100.0	99.5	99.4	99.3
A2, A3, B7, A24, B44, A1, B27, B62, B58	99.9	99.6	100.0	99.8	99.9	99.8

SF 184895 v1

Table XXII. Cross-reactive binding of CEA analog peptides

Source	AA	Sequence	A*0201 nM	A*0202 nM	A*0203 nM	A*0206 nM	A*6802 nM	No. A2 Alleles Bound
CEA.24	9	LLTFWNPPT	179	1720	67	755	-- [2]	2
CEA.24M2V9	9	LMTFWNPPV	4.5	782	7.7	34	3333	3
CEA.24V9	9	LLTFWNPPV	16	307	26	56	952	4
CEA.78	9	QIIGYVIGT	313	148	106	100	150	5
CEA.78L2V9	9	QLIGYVIGV	9.4	5.9	2.3	21	2.3	5
CEA.233	10	VLYGPDAPTI	128	606	270	804	--	2
CEA.233V10	10	VLYGPDAPTV	26	430	16	206	952	4
CEA.411	10	VLYGPDDPTI	294	358	476	7400	--	3
CEA.411V10	10	VLYGPDDPTV	161	105	91	2467	--	3
CEA.569	9	YVCGIQNSV	98	358	159	80	181	5
CEA.569L2	9	YLCGIQNSV	50	24	12	31	3478	4
CEA.589	9	VLYGPDTPI	200	878	53	638	--	2
CEA.589V9	9	VLYGPDTPV	20	165	91	154	9756	4
CEA.605	9	YLSGANLNL	28	165	2.4	804	--	3
CEA.605V9	9	YLSGANLNV	73	13	13	80	1600	4
CEA.687	9	ATVGIMIGV	36	8.8	20	11	0.80	5
CEA.687L2	9	ALVGIMIGV	10	63	31	100	102	5
CEA.691	9	IMIGVLVGV	69	62	13	106	89	5
CEA.691L2	9	ILIGVLVGV	22	8.0	3.2	16	160	5

1) Wild-type peptides presented for reference purposes.
2) -- indicates binding affinity =10,000nM.

TABLE XXII A A01 Analog P ptides

Peptide	AA	Sequence	Source	A*0101 nM
52.0105	11	RVDGNRQIIGY	CEA.72	294.1
52.0109	11	RSDSVILNVLY	CEA.225	47.2
52.0113	11	HSDPVILNVLY	CEA.403	25.8
52.0116	11	RSDPVTLDVLY	CEA.581	7.8
57.0004	9	QQDTPGPAY	CEA.87.D3	56.8
57.0007	9	AADNPPAQY	CEA.261.D3	45.5
57.0008	9	ITDNNSGSY	CEA.289.D3	96.2
57.001	9	VTDNDVGPY	CEA.383.D3	4.1
57.0011	9	PTDSPSYTY	CEA.418.D3	37.9
57.0012	9	TIDPSYTYY	CEA.419.D3	3.1
57.0013	9	AADNPPAQY	CEA.439.D3	44.6
57.0014	9	ITDKNSGLY	CEA.467.D3	11.9
57.0103	10	PTDSPLNTSY	CEA.240.D3	266
57.0104	10	PTDSPSYTYY	CEA.418.D3	1.1
57.0105	10	HTASNPSPQY	CEA.616.T2	131.6
57.0106	10	HSDSNPSPQY	CEA.616.D3	44.6

Table XXII B A03 Analog Peptides

Peptide	AA	Sequence	Source	A*0301 nM	A*1101 nM	A*3101 nM	A*3301 nM	A*6801 nM	A3 XRN
1371.01	10	TVSPLNTSYR	CEA.241.V2	458.3	54.5	187.5	557.7	8.7	4
1371.02	10	TVSPLNTSYK	CEA.241.V2K10	16.9	6.3	10588.2	-48333.3	7.3	3
1371.03	10	RVLTLLSVTR	CEA.376.V2	343.8	222.2	11.3	6041.7	666.7	3
1371.04	10	RVLTLLSVTK	CEA.376.V2K10	37.9	50	163.6	-72500	5714.3	3
1371.05	10	TVSPSYTYYR	CEA.419.V2	2340.4	3000	29	263.6	8.6	3
1371.06	10	TVSPSYTYYK	CEA.419.V2K10	68.8	42.9	3673.5	26363.6	6.7	3
1371.07	9	IVPSYTYYR	CEA.420.V2	91.7	13.3	25.7	58	2.6	5
1371.08	9	IVPSYTYYK	CEA.420.V2K9	17.2	54.5	720	4328.4	21.6	3
1371.09	10	RVLTLFNVTR	CEA.554.V2	297.3	93.8	9	7631.6	42.1	4
1371.1	10	RVLTLFNVTK	CEA.554.V42K10	20.8	31.6	233.8	41428.6	2352.9	3
1371.13	9	FVSNLATGK	CEA.656.K9	1466.7	206.9	-36000	-72500	5.3	2

Table XXIIC A24 Analog P ptides

Peptide	AA	Sequence	Source	A*2401 nM
52.0033	8	IYPNASLL	CEA.101	176.5
52.0038	8	SWFVNGTF	CEA.270	480
52.0137	11	RWCIPWQRLLL	CEA.10	151.9
52.0138	11	PWQRLLLTASL	CEA.14	324.3
52.0141	11	FYTLHVIKSDL	CEA.119	480
52.0142	11	TYLWWVNNQSL	CEA.175	85.7
52.0144	11	TYLWWVNNQSL	CEA.353	46.2
52.0145	11	SYTYYRPGVNL	CEA.423	218.2
52.0146	11	TYYRPGVNLSL	CEA.425	131.9
52.0147	11	TYLWWVNGQSL	CEA.531	92.3
57.0036	9	RYCIPWQRF	CEA.10.Y2F9	190.5
57.0037	9	IYPNASLLF	CEA.101.F9	2.2
57.0038	9	LYWVNNQSF	CEA.177.Y2F9	63.2
57.0039	9	LYGPDAPTF	CEA.234.F9	63.2
57.0041	9	TYYRPGVNF	CEA.425.F9	52.2
57.0042	9	LYWVNGQSF	CEA.533.Y2F9	15.8
57.0044	9	QYSWRINGF	CEA.624.F9	109.1
57.0045	9	TYACFVSNF	CEA.652.F9	8.6
57.0072	10	RYCIPWQRLF	CEA.10.Y2F10	26.1
57.0073	10	FYNPPTTAKF	CEA.27.Y2F10	181.8
57.0074	10	VYPELPKPSF	CEA.140.F10	106.2
57.0075	10	TYQQSTQELF	CEA.276.Y2	307.7
57.0076	10	VYAEPPKPFF	CEA.318.F10	26.7
57.0077	10	YYRPGVNLSF	CEA.426.F10	10
57.0078	10	QYSWLIDGNF	CEA.446.F10	60
57.0079	10	SYLSGANLNF	CEA.604.F10	10

Table XXIII. Immunogenicity of A2 supermotif-bearing peptides

Peptide	AA	Sequence	A*0201 nM	A*0202 nM	A*0203 nM	A*0206 nM	A*6802 nM	No. A2 Alleles Bound	CTL Peptide [1]	CTL Wild-type	CTL Tumor
CEA.78	9	QILGYVIGT	313	148	106	100	151	5		0/3	
CEA.354	10	YLWWVNNQSL	26	108	26	487	333	5		1/2	0/1
CEA.569	9	YVCGIQNSV	98	358	159	80	182	5		1/2	0/1
CEA.605	9	YLSGANLNL	28	165	2	804	-- [2]	3		2/2	1/2
CEA.687	9	ATVGIMIGV	36	9	20	11	1	5		1/1	1/1
CEA.691	9	IMIGVLVGV	69	62	13	106	89	5		8/8	4/7
CEA.24	9	LLTFWNPPT	179	1720	67	755	-- [2]	2		0/1	0/1
CEA.24V9	9	LLTFWNPPV	16	307	26	56	952	4	1/1		1/1
CEA.233	10	VLYGPDAPTI	128	606	270	804	--	2		2/4	0/3
CEA.233V10	10	VLYGPDAPTV	26	430	16	206	952	4	3/4	2/2	1/4
CEA.589	9	VLYGPDTPI	200	878	53	638	--	2		1/1	0/1
CEA.589V9	9	VLYGPDTPV	20	165	91	154	9756	4	2/2	2/2	0/2
CEA.605	9	YLSGANLNL	28	165	2.4	804	--	3		2/2	1/2
CEA.605V9	9	YLSGANLNV	73	13	13	80	1600	4	4/4	3/4	1/4

1) Number of donors yielding a positive response/total tested.
2) -- indicates binding affinity =10,000nM.

Table XXIV. MHC-peptide binding assays: cell lines and radiolabeled ligands.

A. Class I binding assays

Species	Antigen	Allele	Cell line	Radiolabeled peptide	
				Source	Sequence
Human	A1	A*0101	Steinlin	Hu. J chain 102-110	YTAVVPLVY
	A2	A*0201	JY	HBVc 18-27 F6->Y	FLPSDYFPSV
	A2	A*0202	P815 (transfected)	HBVc 18-27 F6->Y	FLPSDYFPSV
	A2	A*0203	FUN	HBVc 18-27 F6->Y	FLPSDYFPSV
	A2	A*0206	CLA	HBVc 18-27 F6->Y	FLPSDYFPSV
	A2	A*0207	721.221 (transfected)	HBVc 18-27 F6->Y	FLPSDYFPSV
	A3		GM3107	non-natural (A3CON1)	KVFPYALINK
	A11		BVR	non-natural (A3CON1)	KVFPYALINK
	A24	A*2402	KAS116	non-natural (A24CON1)	AYIDNYNKF
	A31	A*3101	SPACH	non-natural (A3CON1)	KVFPYALINK
	A33	A*3301	LWAGS	non-natural (A3CON1)	KVFPYALINK
	A28/68	A*6801	C1R	HBVc 141-151 T7->Y	STLPETYVVRR
	A28/68	A*6802	AMAI	HBV pol 646-654 C4->A	FTQAGYPAL
	B7	B*0702	GM3107	A2 sigal seq. 5-13 (L7->Y)	APRTLVYLL
	B8	B*0801	Steinlin	HIVgp 586-593 Y1->F, Q5->Y	FLKDYQLL
	B27	B*2705	LG2	R 60s	FRYNGLIHR
	B35	B*3501	C1R, BVR	non-natural (B35CON2)	FPFKYAAAF
	B35	B*3502	TISI	non-natural (B35CON2)	FPFKYAAAF
	B35	B*3503	EHM	non-natural (B35CON2)	FPFKYAAAF
	B44	B*4403	PITOUT	EF-1 G6->Y	AEMGKYSFY
	B51		KAS116	non-natural (B35CON2)	FPFKYAAAF
	B53	B*5301	AMAI	non-natural (B35CON2)	FPFKYAAAF
	B54	B*5401	KT3	non-natural (B35CON2)	FPFKYAAAF
	Cw4	Cw*0401	C1R	non-natural (C4CON1)	QYDDAVYKL
	Cw6	Cw*0602	721.221 transfected	non-natural (C6CON1)	YRHDGGNVL
	Cw7	Cw*0702	721.221 transfected	non-natural (C6CON1)	YRHDGGNVL
Mouse	D^b		EL4	Adenovirus E1A P7->Y	SGPSNTYPEI
	K^b		EL4	VSV NP 52-59	RGYVFQGL
	D^d		P815	HIV-IIIB ENV G4->Y	RGPYRAFVTI
	K^d		P815	non-natural (KdCON1)	KFNPMKTYI
	L^d		P815	HBVs 28-39	IPQSLDSYWTSL

B. Class II binding assays

Species	Antigen	Allele	Cell line	Radiolabeled peptide	
				Source	Sequence
Human	DR1	DRB1*0101	LG2	HA Y307-319	YPKYVKQNTLKLAT
	DR2	DRB1*1501	L466.1	MBP 88-102Y	VVHFFKNIVTPRTPPY
	DR2	DRB1*1601	L242.5	non-natural (760.16)	YAAFAAAKTAAAFA
	DR3	DRB1*0301	MAT	MT 65kD Y3-13	YKTIAFDEEARR
	DR4w4	DRB1*0401	Preiss	non-natural (717.01)	YARFQSQTTLKQKT
	DR4w10	DRB1*0402	YAR	non-natural (717.10)	YARFQRQTTLKAAA
	DR4w14	DRB1*0404	BIN 40	non-natural (717.01)	YARFQSQTTLKQKT
	DR4w15	DRB1*0405	KT3	non-natural (717.01)	YARFQSQTTLKQKT
	DR7	DRB1*0701	Pitout	Tet. tox. 830-843	QYIKANSKFIGITE
	DR8	DRB1*0802	OLL	Tet. tox. 830-843	QYIKANSKFIGITE
	DR8	DRB1*0803	LUY	Tet. tox. 830-843	QYIKANSKFIGITE
	DR9	DRB1*0901	HID	Tet. tox. 830-843	QYIKANSKFIGITE
	DR11	DRB1*1101	Sweig	Tet. tox. 830-843	QYIKANSKFIGITE
	DR12	DRB1*1201	Herluf	unknown eluted peptide	EALIHQLKINPYVLS
	DR13	DRB1*1302	H0301	Tet. tox. 830-843 S->A	QYIKANAKFIGITE
	DR51	DRB5*0101	GM3107 or L416.3	Tet. tox. 830-843	QYIKANAKFIGITE
	DR51	DRB5*0201	L255.1	HA 307-319	PKYVKQNTLKLAT
	DR52	DRB3*0101	MAT	Tet. tox. 830-843	NGQIGNDPNRDIL
	DR53	DRB4*0101	L257.6	non-natural (717.01)	YARFQSQTTLKQKT
	DQ3.1	A1*0301/DQB1*0	PF	non-natural (ROIV)	YAHAAHAAHAAHAAHAA
Mouse	IA^b		DB27.4	non-natural (ROIV)	YAHAAHAAHAAHAAHAA
	IA^d		A20	non-natural (ROIV)	YAHAAHAAHAAHAAHAA
	IA^k		CH-12	HEL 46-61	YNTDGSTDYGILQINSR
	IA^s		LS102.9	non-natural (ROIV)	YAHAAHAAHAAHAAHAA
	IA^u		91.7	non-natural (ROIV)	YAHAAHAAHAAHAAHAA
	IE^d		A20	Lambda repressor 12-26	YLEDARRKKAIYEKKK
	IE^k		CH-12	Lambda repressor 12-26	YLEDARRKKAIYEKKK

Table XXV. Antibodies used in MHC purification.

Monoclonal antibody	Specificity
W6/32	HLA-class I
B123.2	HLA-B and C
IVD12	HLA-DQ
LB3.1	HLA-DR
M1/42	H-2 class I
28-14-8S	H-2 Db and Ld
34-5-8S	H-2 Dd
B8-24-3	H-2 Kb
SF1-1.1.1	H-2 Kd
Y-3	H-2 Kb
10.3.6	H-2 IAk
14.4.4	H-2 IEd, IEK
MKD6	H-2 IAd
Y3JP	H-2 IAb, IAs, IAu

Table XXVI. Crossbinding data of A2 supermotif peptides.

Source	AA	Sequence	A*0201 nM	A*0202 nM	A*0203 nM	A*0206 nM	A*6802 nM	No. A2 Alleles Crossbound
CEA.24	9	LLTFWNPPT	179	1720	67	755	--	2
CEA.78	9	QIIGYVIGT	313	148	106	100	150	5
CEA.233	10	VLYGPDAPTI	128	606	270	804	--	2
CEA.354	10	YLWWVNNQSL	26	108	26	487	67	5
CEA.411	10	VLYGPDDPTI	294	358	476	7400	--	3
CEA.432	9	NLSLSCHAA	455	2867	1449	18500	--	1
CEA.532	10	YLWWVNGQSL	33	331	21	2056	286	4
CEA.569	9	YVCGIQNSV	98	358	159	80	181	5
CEA.589	9	VLYGPDTPI	200	878	53	638	--	2
CEA.605	9	YLSGANLNL	28	165	2.4	804	--	3
CEA.687	9	ATVGIMIGV	36	8.8	20	11	0.80	5
CEA.690	10	GIMIGVLVGV	64	205	31	142	500	5
CEA.691	9	IMIGVLVGV	69	62	13	106	89	5
CEA.691	10	IMIGVLVGVA	227	68	44	726	1509	3

-- indicates binding affinity =10,000nM.

Table XXVII. Immunogenicity of A2 supermotif peptides

Source	AA	Sequence	A*0201 nM	A*0202 nM	A*0203 nM	A*0206 nM	A*6802 nM	No. A2 Alleles Crossbound	CTL Wild-type [1]	CTL Tumor
CEA.78	9	QIIGYVIGT	313	148	106	100	151	5	0/3	
CEA.354	10	YLWWVNNQ	26	108	26	487	333	5	1/2	0/1
CEA.569	9	YVCGIQNSV	98	358	159	80	182	5	1/2	0/1
CEA.605	9	YLSGANLNL	28	165	2.4	804	-- [2]	3	2/2	1/2
CEA.687	9	ATVGIMIGV	36	8.8	20	11	0.80	5	1/1	1/1
CEA.691	9	IMIGVLVGV	69	62	13	106	89	5	8/8	4/7

1) Number of donors yielding a positive response/total tested.
2) -- indicates binding affinity =10,000nM.

Table XXVIII. Immunogenicity A2 supermotif analog peptides

Source	AA	Sequence	A*0201 nM	A*0202 nM	A*0203 nM	A*0206 nM	A*6802 nM	No. A2 Alleles Crossbound	CTL Peptide [1]	CTL Wild-type	CTL Tumor
CEA.24	9	LLTFWNPPT	179	1720	67	755	-- [2]	2		0/1	0/1
CEA.24V9	9	LLTFWNPPV	16	307	26	56	952	4	1/1		1/1
CEA.233	10	VLYGPDAPTI	128	606	270	804	--	2		2/4	0/3
CEA.233V10	10	VLYGPDAPTV	26	430	16	206	952	4	3/4	2/2	1/4
CEA.589	9	VLYGPDTPI	200	878	53	638	--	2		1/1	0/1
CEA.589V9	9	VLYGPDTPV	20	165	91	154	9756	4	2/2	2/2	0/2
CEA.605	9	YLSGANLNL	28	165	2.4	804	--	3		2/2	1/2
CEA.605V9	9	YLSGANLNV	73	13	13	80	1600	4	4/4	3/4	1/4

1) Number of donors yielding a positive response/total tested.
2) -- indicates binding affinity =10,000nM.

Table XXIX. DR supertype primary binding

Peptide	DR147 Algo Sum	Sequence	Source	DR1 nM	DR4w4 nM	DR7 nM	DR147 Cross-reactivity
39.0217	2	RWCIPWQRLLLTASL	CEA.10	8.2	542	357	3
39.0218	3	QRLLLTASLLTFWNP	CEA.16	--	--	--	0
39.0219	2	EVLLLVHNLPQHLFG	CEA.50	2.0	52	53	3
39.0220	3	GREIIYPNASLLIQN	CEA.97	8.1	484	45	3
39.0221	2	EIIYPNASLLIQNII	CEA.99	14	1154	156	2
39.0222	2	NASLLIQNIIQNDTG	CEA.104	4546	--	--	0
39.0223	3	DTGFYTLHVIKSDLV	CEA.116	69	1731	227	2
39.0224	2	YPELPKPSISSNNSK	CEA.141	5556	--	--	0
39.0225	2	KPSISSNNSKPVEDK	CEA.146	2381	--	7576	0
39.0226	3	YLWWVNNQSLPVSPR	CEA.176	0.59	8.0	42	3
39.0227	3	LWWVNNQSLPVSPRL	CEA.177	217	1552	3049	1
39.0228	2	QYSWFVNGTFQQSTQ	CEA.268	192	80	926	3
39.0229	2	DTGLNRTTVTTITVY	CEA.305	--	--	2841	0
39.0230	2	KPFITSNNSNPVEDE	CEA.324	--	--	--	0
39.0231	2	NRTLTLLSVTRNDVG	CEA.375	238	--	--	1
39.0232	2	QELFISNITEKNSGL	CEA.460	--	2500	--	0
39.0233	3	RTTVKTITVSAELPK	CEA.488	455	7031	317	2
39.0234	2	SAELPKPSISSNNSK	CEA.497	--	--	--	0
39.0235	2	LDVLYGPDTPIISPP	CEA.587	--	--	--	0
39.0236	2	TQVLFIAKITPNNNG	CEA.637	61	--	6579	1
39.0237	2	QVLFIAKITPNNNGT	CEA.638	42	1875	--	1
39.0238	3	YACFVSNLATGRNNS	CEA.653	208	1667	3571	1
39.0239	2	NNSIVKSITVSASGT	CEA.665	91	25	676	3
39.0240	3	NSIVKSITVSASGTS	CEA.666	78	25	329	3

-- indicates binding affinity =10,000nM.

Table XXX DR supertype crossbinding

Peptide	Sequence	Source	DR1 nM	DR4w4 nM	DR7 nM	DR2w2 ß1 nM	DR2w2 ß2 nM	DR6w1 9 nM	DR5w1 1 nM	DR8w2 nM	DR147 Degen	Broad Degen (5/8)
39.0217	RWCIPWQRLLLTASL	CEA.10	8.2	542	357	827	--	318	--	--	3	5
39.0219	EVLLLVHNLPQHLFG	CEA.50	2.0	52	53	40	--	1.0	588	408	3	7
39.0220	GREIIYPNASLLIQN	CEA.97	8.1	484	45	24	8333	2.9	6897	5904	3	5
39.0221	EIIYPNASLLIQNII	CEA.99	14	1154	156	57	--	11	--	--	2	4
39.0223	DTGFYTLHVIKSDLV	CEA.116	69	1731	227	506	800	3889	2500	790	2	5
39.0226	YLWWVNNQSLPVSPR	CEA.176	0.60	8.0	42	110	2105	2.3	29	1065	3	6
39.0228	QYSWFVNGTFQQSTQ	CEA.268	192	80	926	--	6061	5833	370	--	3	4
39.0233	RTTVKTITVSAELPK	CEA.488	455	7031	317	364	--	700	--	--	2	4
39.0239	NNSIVKSITVSASGT	CEA.665	91	25	676	3138	--	51	--	4083	3	4
39.0240	NSIVKSITVSASGTS	CEA.666	78	25	329	3957	--	76	--	2882	3	4

-- indicates binding affinity =10,000nM.

Table XXXI. DR3 binding

Peptide	Sequence	Source	DR3 nM
39.0313	QNIIQNDTGFYTLHV	CEA.110	938
39.0314	LHVIKSDLVNEEATG	CEA.122	2308
39.0315	KSDLVNEEATGQFRV	CEA.126	--
39.0316	SDLVNEEATGQFRVY	CEA.127	--
39.0317	NEEATGQFRVYPELP	CEA.131	--
39.0318	QFRVYPELPKPSISS	CEA.137	--
39.0319	AVAFTCEPETQDATY	CEA.162	--
39.0320	TASYKCETQNPVSAR	CEA.210	--
39.0321	NVLYGPDAPTISPLN	CEA.232	--
39.0322	ISPLNTSYRSGENLN	CEA.242	--
39.0323	SGSYTCQAHNSDTGL	CEA.294	--
39.0324	TITVYAEPPKPFITS	CEA.315	--
39.0325	SNPVEDEDAVALTCE	CEA.332	--
39.0326	AVALTCEPEIQNTTY	CEA.340	--
39.0327	NQSLPVSPRLQLSND	CEA.360	--
39.0328	RLQLSNDNRTLTLLS	CEA.368	938
39.0329	ECGIQNELSVDHSDP	CEA.392	--
39.0330	QNELSVDHSDPVILN	CEA.396	3659
39.0331	NVLYGPDDPTISPSY	CEA.410	--
39.0332	GVNLSLSCHAASNPP	CEA.430	--
39.0333	TITVSAELPKPSISS	CEA.493	--
39.0334	AVAFTCEPEAQNTTY	CEA.518	--
39.0335	SDPVTLDVLYGPDTP	CEA.582	--
39.0336	DVLYGPDTPIISPPD	CEA.588	--
39.0337	GANLNLSCHSASNPS	CEA.608	--

-- indicates binding affinity =10,000nM.

Table XXXII. HTLCandidate Epitopes

Peptide	Sequence	Motif	Source	DR1 nM	DR4w4 nM	DR7 nM	DR3 nM	DR2w2 ß1 nM	DR2w2 ß2 nM	DR6w1 9 nM	DR5w1 1 nM	DR8w2 nM	DR147 Cross-reactivity	Broad Cross-reactivity (5/8)	DR3 Binder
39.0217	RWCIPWQRLLLTASL	DR sup	CEA.10	8.2	542	357	--	827	--	318	--	--	3	5	0
39.0219	EVLLLVHNLPQHLFG	DR sup	CEA.50	2.0	52	53	336	40	--	1.0	588	408	3	7	1
39.0220	GREIIYPNASLLIQN	DR sup	CEA.97	8.1	484	45	1123	24	8333	2.9	6897	5904	3	5	0
39.0313	QNIIQNDTGFYTLHV	DR3	CEA.110	1136	>8182	--	938	867	--	9.7	--	--	0	2	1
39.0223	DTGFYTLHVIKSDLV	DR sup	CEA.116	69	1731	227	--	506	800	3889	2500	790	2	5	0
39.0226	YLWWVNNQSLPVSPR	DR sup	CEA.176	0.60	8.0	42	2310	110	2105	2.3	29	1065	3	6	0
39.0328	RLQLSNDNRTLTLLS	DR3	CEA.368	--	>8182	--	938	--	--	729	--	--	0	1	1

-- indicates binding affinity =10,000nM.